Universität Zürich[UZH]

THE PERISYLVIAN BRAIN OF ABSOLUTE PITCH MUSICIANS

Thesis (cumulative thesis)

presented to the Faculty of Arts and Social Sciences

of the University of Zurich

for the degree of Doctor of Philosophy

by

Anja Burkhard

Accepted in the spring semester 2019

on the recommendation of the doctoral committee composed of

Prof. Dr. rer. nat. Lutz Jäncke (main supervisor)

Prof. Dr. Sascha Frühholz

Zurich, 2020

CONTENTS

Acknowledgments 1

Abstract 3

Zusammenfassung 5

1. General Introduction 9

1.1 Short history of cognitive neuroscience and plasticity 9

1.2 Musicians as model for neuroplasticity 10

1.3 Absolute pitch and relative pitch 12

1.3.1 The psychology of absolute pitch 13

1.3.2 The neuroscientific foundations of absolute pitch 15

1.4 Open questions 16

1.5 Methods 18

1.5.2 The origin of the EEG signal and source reconstruction 19

1.5.1 Event-related potentials 21

1.6 Methods: Diffusion tensor imaging 26

1.6.1 Tract-based spatial statistics 29

2. Empirical Part Study I 31

2.1 Abstract 32

2.2 Introduction 32

2.3 Materials and Methods 37

2.3.1 Subjects 37

2.3.2 AP Test 40

2.3.3 Task 41

2.3.4 Stimuli 42

2.3.5 EEG recording and pre-processing 43
2.3.6 ERP analyses 44
2.3.7 Source reconstruction 45
2.3.8 Statistical analysis 48
2.4 Results 50
2.4.1 Results with respect to the amplitudes of the N100, P200, N200 and N400 components 50
2.4.2 Results with respect to the intracortical sources of the N100 and P200 components 51
2.5 Discussion 53
2.5.1 General discussion 53
2.5.2 Discussion with respect to the ERP components 54
2.5.3 Discussion on the intracortical sources of the N100 component 57
2.5.4 Possible implications for AP perception and labelling 60
2.6 Limitations and Outlook 61
2.7 Conclusion 62
3. Empirical Part Study II 63
3.1 Abstract 64
3.2 Keywords 64
3.3 Introduction 64
3.4 Materials and Methods 66
3.4.1 participants 66
3.4.2 DTI data acquisition and preprocessing 68
3.5 Results 72

3.5.1 Behavioral data 72
3.5.2 Results of the fa analyses 73
3.6 Discussion 75
3.6.1 future perspectives 78
3.6.2 Limitations 80
3.7 Conclusions 81
4. General discussion 83
4.1 AP psychological aspects 85
4.1.1 Absolute pitch and categorization and labeling 85
4.1.2 Absolute pitch and attention 87
4.2 Methodological considerations 88
4.2.1 Absolute pitch classification 89
4.2.2 Event-related potential technique: signed area 91
4.2.3 Functional and structural conformities 93
4.3 Future perspectives 95
4.4 Conclusion 97
References 99
List of Figures 120
List of Tables 121
Appendix 122
A Supplementary Material 122
B Curriculum Vitae 123

ACKNOWLEDGMENTS

The support of so many people made this work possible, and an adequate expression of thanks to all of you would go beyond the scope of this acknowledgment paragraph because it would definitively surpass the page limits. Therefore, my gratitude goes first to all the people I know for their continuous support and assistance at all levels. Thank you!

I am expressing my deepest gratitude to Prof. Dr. Lutz Jäncke. This applies not only to his scientific support, his motivating words and likable manner, and his openness to any type of requests; but most importantly, to his being able to arouse my curiosity about the human brain. At the early beginning of my studies, I could not imagine that I would follow the path of neuroscience. It was a gift to attend lectures about the human brain by Prof. Dr. Jäncke with his enthusiastic character and plenty of humor that relieved many situations even in my time as a PhD student. Thank you!

Special thanks go to Stefan Elmer and Jürgen Hänggi for the mentoring, scientific guidance, and many pleasant conversations about the entire world. My sincere gratitude also goes to Prof. Sascha Frühholz for joining my doctoral committee and Dr. Jonny Draffin for proofreading and linguistic revision of this thesis. Furthermore, I would like to thank Marielle, Simon, and especially my roommate Christian for realizing the absolute pitch project together with me, always with great confidence and without losing sense of humor. My thanks also go to my colleagues at the department of neuropsychology for the warm and friendly working atmosphere, the stimulating conversations, and the many good moments we shared. Thank you Yolanda, Carina, Lisa, Seyed, Michael, Matthias, Mrs. Jäncke, and Silvano. Also, big thanks to Michaela, our good soul at the office, for the many good talks and taking care of all that nasty administrational stuff. Many thanks to you, Martin Meyer, for your thoughtful inputs and the humorous conversations, particularly about our shared favorite topic: dogs.

Big thanks are also extended to all those who participated in our studies. All this would not have been possible without you! You turned the monotonous measurement routine into a pleasant experience. I am very happy that I had the pleasure to get to know you all.

My deepest thanks go to my parents. In particular, I would like to take the opportunity to thank you for never telling me: “You cannot do this, it is not possible”. This gave me the mindset that everything is possible and this is an important source of strength in anything I am doing. Thank you! The same holds true for my companion Max von Gellhorn and my very best friends Martin Elsener and Steffen Tepel. Thank you for your warm-hearted encouragement and ongoing support. Thank you!

This work was supported by the Swiss National Foundation (SNSF).

ABSTRACT

Absolute pitch (AP) is an ability that very few musicians are gifted with. Prevalence within the population of musicians has been estimated to 10–15% (Baharloo et al., 1998; Gregersen et al., 2001; Gregersen et al., 1999; Hirose et al., 2002). In contrast to relative pitch, which is a common ability among musicians, AP enables its possessor to identify or produce a specific tone without the use of a reference tone. The neural mechanisms underlying AP are not yet fully understood. Nonetheless, numerous studies indicate that this extraordinary ability is accompanied by special functional as well as structural features. Given the rare occurrence of AP, experimental studies with large sample sizes are lacking. This dissertation consists of two studies, which aim to investigate the neuronal mechanisms underlying AP by examining a relatively large sample of absolute and relative pitch possessors.

Study I was conducted by using electroencephalography and focuses on the functional aspects of AP. The critical time frames concerning neural processes essential for AP were examined based on two prevailing theories. The first theory postulates that early processes such as auditory perception and encoding are pivotal (Hirata et al., 1999; Schulze et al., 2009; Wengenroth et al., 2014). The second so-called two-component model states that a later occurring process dubbed “labeling” is essential for AP (Levitin & Rogers, 2005). While the labeling process plays an important role for AP in the aforementioned model, pitch memory is considered to be widely distributed among the common population and therefore as being negligible for AP. With an analysis of the event-related potentials N100, P200, N200, and N400, we show that AP possessors differ from musicians with relative pitch mainly in the early N100 component. This difference occurred during active listening to sine tones. Furthermore, performing source estimation revealed that the activation underlying the N100 component varies mostly in the perisylvian cortex (superior temporal gyrus, middle temporal gyrus and heschl’s gyrus) between the two groups of musicians.

Study II, on the other hand, examines the structural aspects of the AP ability by using diffusion tensor magnetic resonance imaging. This study focuses on perisylvian brain areas, which are related to the perception, encoding, and categorization of auditory information. An arbitrary group classification was avoided by using a correlational approach. In this approach, the AP ability is treated as a continuous measure. The individual scores of an AP test indicated the proficiency level of AP and were correlated with fractional anisotropy values of perisylvian brain regions. The perisylvian regions were selected following an analysis of the existing literature. The analysis revealed enhanced fractional anisotropy values with increasing AP proficiency levels in the white matter underlying the left-sided planum temporale.

In summary, we show here that AP possessors differed from their colleagues with relative pitch in the very early stages of auditory processing (≈100 ms). Consistent with this finding, white matter features of a perisylvian brain region known to be important for AP correlated with AP proficiency level. This brain region is involved in categorization and early auditory processing steps. However, no supporting evidence for automatic labeling, as stated in the two-component model, was found in AP possessors. Thus, these results point to early auditory processing differences at the perception and categorization levels, dependent on the proficiency level of AP. Future studies should differentiate the specific roles of the involved perisylvian brain areas regarding their functional and structural characteristics.

ZUSAMMENFASSUNG

Das absolute Gehör ist eine Fähigkeit, die nur wenigen Musikern vorbehalten ist. So wurde die Prävalenz der Absolut-Hörer innerhalb der Musiker-Population auf 10-15% geschätzt (Baharloo et al., 1998; Gregersen et al., 2001; Gregersen et al., 1999; Hirose et al., 2002). Im Gegensatz zum relativen Gehör, über das die meisten guten Musiker verfügen, befähigt das absolute Gehör den Musiker dazu, einen Ton ohne Hilfe eines Referenz-Tons zu benennen oder aber nach vorgegebener Tonbezeichnung korrekt zu produzieren. Über die dem absoluten Gehör zugrunde liegenden neuronalen Mechanismen ist man sich noch nicht vollends im Klaren. Etliche Studien konnten jedoch bisher belegen, dass diese spezielle Fähigkeit mit funktionellen, wie auch strukturellen Besonderheiten einhergeht. Da das absolute Gehör nur sehr selten auftritt, fehlen jedoch dementsprechend experimentelle Studien mit grossen Versuchspersonen-Zahlen. Die vorliegende Dissertation setzt sich aus zwei Studien zusammen, deren Ziel es ist, anhand einer relativ grossen Stichprobe von Absolut- und Relativ-Hörern, den dem Absoluten Gehör zugrunde liegenden neuronalen Mechanismen auf den Grund zu gehen.

Studie I wurde mittels Elektroenzephalographie durchgeführt und konzentriert sich auf die funktionellen Grundlagen des absoluten Gehörs. Ausgehend von den zwei vorherrschenden Theorien wurde das zeitliche Auftreten der für das absolute Gehör essentiellen Verarbeitungs-Prozesse untersucht. Die erste Theorie geht davon aus, dass früh stattfindende, auditorische Wahrnehmungs- und Enkodierungs-Prozesse ausschlaggebend sind für das absolute Gehör (Hirata et al., 1999; Schulze et al., 2009; Wengenroth et al., 2014). Die zweite, sogenannte Zwei-Komponenten-Theorie hingegen besagt, dass ein später erfolgender Prozess, dass sogenannte „Labeling“, also die Ton-Benennung, von essentieller Bedeutung für das absolute Gehör ist (Levitin & Rogers, 2005). Im Gegensatz dazu ist die andere Komponente, das Pitch-Gedächtnis, in der Bevölkerung weit verbreitet und somit, zumindest für die Ausprägung des absoluten

Gehörs, nicht von allzu grosser Bedeutung. Mittels einer Analyse der Ereigniskorrelierten Potentiale N100, P200, N200 und N400 konnte gezeigt werden, dass sich Musiker mit absolutem Gehör vor allem in der frühen Komponente N100 von Musikern mit relativem Gehör unterscheiden. Dies während dem aktiven Anhören von Sinustönen. Des Weiteren wurde mittels Quellenschätzung festgestellt, dass sich die der N100 zugrunde liegende Aktivierung vor allem im rechten perisylvischen Kortex (STG, MTG und HG) zwischen den beiden Musikergruppen unterscheidet.

Studie II hingegen untersucht die strukturellen Aspekte des absoluten Gehörs mittels Diffusionstensor-Magnetresonanztomographie. Die Studie fokussiert auf perisylvische Gehirnareale, welche mit der Wahrnehmung, Enkodierung und Kategorisierung tonaler Information in Verbindung gebracht werden. Eine willkürliche Gruppeneinteilung wurde durch einen korrelativen Ansatz vermieden. Durch diesen Ansatz wird die Fähigkeit des absoluten Gehörs als kontinuierlich behandelt. Die individuellen Messwerte eines Tests für das absolute Gehör, welche auf das Fähigkeitsniveau hindeuten, wurden mit den Werten der fraktionalen Anisotropie in perisylvischen Gehirn-Strukturen korreliert. Die perisylvischen Regionen wurden aufgrund der derzeitigen Literatur ausgewählt. Die Analyse ergab erhöhte fraktionale Anisotropie bei zunehmendem Leistungsniveau des absoluten Gehörs in der weissen Substanz unterhalb des linksseitigen Planum temporales.

Zusammenfassend konnte gezeigt werden, dass sich absolut hörende Musiker bereits in sehr frühen Stadien des auditorischen Verarbeitungsprozesses (≈100ms) von ihren Kollegen mit einem relativen Gehör unterscheiden. In Übereinstimmung dazu korrelierten Eigenschaften der weissen Substanz einer perisylvischen Gehirnregion, von welcher man weiss, dass sie für das absolute Gehör wichtig ist, mit dem Leistungsniveau des absoluten Gehörs. Diese Gehirnregion ist in die Kategorisierung und frühe auditorische Verarbeitungsprozesse involviert. Allerdings wurden keine Indizien für eine, gemäss der Zwei-Komponenten-Theorie den absolut hörenden Musikern vorbehaltene, automatische Ton-Benennung gefunden. Somit deuten die

Resultate auf frühe auditorische Verarbeitungs-Unterschiede in Abhängigkeit vom Leistungsniveau des absoluten Gehörs hin, welche auf der Verarbeitungsstufe von Wahrnehmung und Kategorisierung stattfinden. Zukünftige Studien sollten die spezifischen Rollen der beteiligten perisylvischen Hirnareale bezüglich funktionellen und strukturellen Unterschieden ausdifferenzieren.

1. General Introduction

The field of neuropsychology is developing so rapidly that specialization within a sub-area is not only advisable but almost obligatory. The history of cognitive neuroscience is briefly described below, with particular attention to one of the most fascinating properties of our brain: its plasticity. This thesis deals specifically with a group of people who have a very special talent: musicians who have the gift of absolute hearing. The psychological as well as neuroscientific properties of this ability, which were retrospectively ascribed to great artists such as Bach, Beethoven, Mozart, and Jimi Hendrix are described in the following sections. Moreover, open questions regarding our research topic and the methods applied in the two studies will be outlined thoroughly.

1.1 Short history of cognitive neuroscience and plasticity

The investigation of the human brain has been an object of fascination for a long time, albeit with some interruptions. The first known documentation of brain injuries was found in the ancient Edwin Smith papyrus rolls in Egypt, which are supposed to date back to 1,700 BC. The dissection of bodies associated with the interest in the human brain has, for a long time, been a taboo subject, especially during antiquity. The practical and theoretical achievements of the anatomist Andreas Vesalius were groundbreaking. In front of an audience in Bologna, he performed a dissection in which two conditions were notable: first, he criticized the carelessness with which other anatomists performed their craft. Second, he performed the dissection, the presentation, and the explanation of the particular body parts by himself, which was not common practice at that time. Furthermore, in 1543 the anatomic atlas “De humani corporis fabrica” was published in seven volumes by the then 29-year-old Vesalius (Jenny, 2019). At that time, the human brain was considered to be nothing more than a cooling and supporting system for the ventricles, which were thought to be responsible for most

human abilities (Jäncke, 2005). With the advent of phrenology (sometimes referred to as craniology), a view of the brain as the initiator of psychological functions gained importance. Influenced by the localization theory, the main idea was that the expression of certain psychological functions would be visible from the outside in the form of the individual's skull (Strohmer, 2014). Even though the phrenological approach was proven to be erroneous for several reasons, efforts were made to localize psychological functions to specific brain regions. In this context, the clinical case descriptions of Paul Broca (1861) and Carl Wernicke (1874), which describe two different forms of aphasia, should be mentioned (Jäncke, 2005).

Currently, the trend in neuroscientific research is progressing from the classic localization view, in which a specific psychological function is exclusively processed in a single brain region, to a view that emphasizes the network character of the brain. The current perspective combines the idea that several psychological functions are predominantly processed in specific brain regions and above all that an efficiently functioning brain depends on the complex interaction of those regions. A term that is closely related to the network character of the human brain is plasticity. Several decades ago, the prevailing opinion was that an adult brain is unable to restructure itself. This view has now changed, and we know that the human brain is able to alter its anatomical structure up to old age as a consequence of practice and experience (Bennett et al., 1996; Jäncke, 2017; Markham & Greenough, 2004; Sirevaag & Greenough, 1987).

1.2 Musicians as model for neuroplasticity

In recent years, the number of published papers that focus on musicians and music listening in general has increased considerably. Musicality is unique to human beings. Music is a medium that is able to communicate emotions: it synchronizes crowds of people (be it in war or during a concert) and influences moods, to give two examples. However, the source of interest in music or musicians in the context of scientific research is a distinct question. Such fascination, worthy of a mythos, has emerged

around the so-called Mozart effect. In short, spatial abilities are enhanced after listening to a Mozart sonata compared to sitting in silence or listening to relaxing instructions (Rauscher et al., 1993). The intense neuroscientific interest in music is more a consequence of musical training than a direct result of music listening. Generally, professional musicians represent an ideal model for investigating neuroplasticity.

We know from the seminal publications of Donald Hebb that the longer and the more intensively something is practiced, the stronger the accompanying structural changes (Hebb, 1949). Hence, an occasion that permits researchers to investigate plasticity exists in the form of practice and experience, which are major stimulants of plasticity. A suitable model for investigating these practice-induced brain changes are professional musicians. A musician needs to perform several demanding tasks, often simultaneously. A short description of the different tasks that a violinist has to manage can be found in an article from Schlaug (2015). According to the author, the following tasks have to be accomplished and are also practiced while playing the violin:

„ …translating visual analysis of musical notation into motor movements, coordinating multisensory information with bimanual motor activity, developing fine-motor skills mostly of their nondominant hand coupled with metric precision, and monitoring auditory feedback to fine-tune a performance in progress." (Schlaug, 2015).

Furthermore, these abilities are intensively trained over many years and the amount of training can be quantified. A study carried out by Ericsson (1993) impressively demonstrates the effect of musical practice. The authors showed that those musicians that most frequently practiced not only achieved the best graduation results but also had the best prospects of professional success. In other words, even among the best of the best, it is the training effort that is the distinguishing factor.

Within the population of musicians, a subgroup possesses the rare ability to identify, label, and produce the chroma of a specific tone without an internal or external reference (Bachem, 1955; Miyazaki, 1989; Takeuchi & Hulse, 1993). This ability is called absolute pitch (AP) and previous studies estimated the prevalence of 10–15% within the population of professional musicians (Baharloo et al., 1998; Gregersen et al., 2001; Gregersen et al., 1999; Hirose et al., 2002). Remarkably, an absolute pitch possessor (APP) has up to 70 perceptual categories for identifying the pitch of a chroma, whereas other people possess around 6–8 categories. Furthermore, an APP is able to associate a label with each of these categories (Zatorre, 2003). In contrast, the ability of relative pitch (RP) is very common among highly trained musicians. Like most people, relative pitch possessors (RPP) process tones relatively and not absolutely (Takeuchi & Hulse, 1993), as APPs do. RP enables the identification of tone intervals (McDermott & Oxenham, 2008) and the identification of the chroma of a specific tone if a reference tone is given by calculating the difference between the reference and target tones.

However, the neurophysiological and cognitive underpinnings of the processes that enable AP are not entirely understood. It is still unclear to what degree environmental factors or hereditary transmission account for the development of AP (Levitin & Rogers, 2005; Levitin, 1994; Zatorre, 2003). There is also some disagreement in the existing literature with respect to the time frame and its related processing stages during which the AP-specific processing steps occur. Moreover, there is not full conformity with respect to the brain areas involved and the corresponding functions enabling AP. The following sections will summarize what is known about these open questions.

1.3.1 THE PSYCHOLOGY OF ABSOLUTE PITCH

It is important to note that an APP is able to effortlessly identify or produce the chroma of a given pitch. Potentially, tone labeling may be initiated automatically and an APP may not be able to exert control over it (Levitin & Rogers, 2005). Even the development of AP itself seems to require no special efforts (Bachem, 1940; Petran, 1932; Profita & Bidder, 1988). Moreover, all attempts to acquire AP as an adult have failed, or at least lead to an imperfect form of AP (Takeuchi & Hulse, 1993). Meanwhile, it is generally assumed that a specific interplay between environmental factors and cognitive and biological dispositions is needed to develop AP (Bermudez & Zatorre, 2009). It has been hypothesized that musical training must occur before a certain age (9–12 years; Zatorre, 2003) and that genetics may also play a role. The higher incidence of AP in identical twins and also in people of Asian descent speaks in favor of a genetic component, even if such an effect has not yet been proven (Deutsch, 2002; Deutsch et al., 2009).

However, another debate centers on the question of which perceptual and/or cognitive processes (and corresponding brain areas) enable some musicians to hear pitch absolutely. The theory that probably produces the most resonance is the two-component model (Levitin, 1994). As the name implies, the author assumes that AP consists of two different cognitive processes, namely pitch memory and pitch labeling. The former, pitch memory, describes a long-term representation of pitches in memory and constitutes a necessary requirement for AP. However, many studies have shown that stable long-term memory for pitch is widespread in the entire population, even among non-musicians (Schellenberg & Trehub, 2003; Terhardt & Dixon Ward, 1982; Van Hedger et al., 2017). Therefore, pitch memory does not appear to be the pivotal factor that ultimately determines whether or not a person develops AP ability. According to the two-component model, the critical factor is whether one possesses (more or less) stable links between the aforementioned long-term representations of pitch and nominal labels (verbal or otherwise). This describes the second component of the model: the so-called pitch-labeling. An alternative view focuses in particular on early pitch processing

and has been suggested and examined by many previous studies (Hirata et al., 1999; Schulze et al., 2009; Wengenroth et al., 2014). In general, it is assumed that the differences between AP and RP musicians lie at the precognitive level of pitch perception, which predominantly occurs in perisylvian brain areas.

To find evidence that adjudicates among these theories, the classification of AP must first be clarified. Therefore, it is helpful to take a closer look at the distribution and the evaluation of this very rare ability. Many authors have argued that AP ability follows a bipolar distribution (Athos et al., 2007; Carroll, 1975; Miyazaki, 1988). In contrast, other authors have suggested that AP ability is instead continuously distributed and therefore does not necessarily occur in an all-or-nothing manner (Bermudez & Zatorre, 2009; Itoh et al., 2005; Levitin & Rogers, 2005; Wilson et al., 2009). For example, Bachem (1937) assumed in an early study that some musicians possess a single internal reference tone (e.g. the deepest tone one is able to sing) from which all other tones can be inferred. Furthermore, the occurrence of a partial AP ability is more frequently found than a perfect or absent AP ability (Wengenroth et al., 2014).

One of the most important criteria for explaining these partially contrasting views on the distribution pattern of AP ability is probably the various manners in which the ability itself is surveyed. Disagreement prevails even over the selection of parameters thought to be the most important for an AP classification (e.g. reaction times). In most cases, the AP tests require the identification of specific tones. This tone identification has to be accomplished without a known reference tone; otherwise, RP ability would be surveyed. Apart from this generally accepted basic framework, the single components can vary greatly. Features such as the stimulus itself (e.g. piano, sine, or multi-harmonic tones), stimulus length, amount of trials, response form (e.g. verbal or written), and also analysis parameters (e.g. selected features for analysis, scoring systems, etc.) chosen by the individual researcher, are considered as such greatly differing components, to name just a few. This topic will be discussed in detail in Section 4 of this doctoral thesis (general discussion).

1.3.2 THE NEUROSCIENTIFIC FOUNDATIONS OF ABSOLUTE PITCH

As described in the previous chapter, some studies have pointed to early perceptual processes as being the important ones (Burkhard et al., 2019; Hirata et al., 1999; Schulze et al., 2009; Wengenroth et al., 2014), while others have assumed that the later occurring stages weighted toward cognitive processing determine AP (Levitin & Rogers, 2005; Levitin, 1994; Zatorre, 2003). In general, the latter are thought to be primarily modulated by frontal brain regions, whereas the former are assumed to be generated in perisylvian brain regions. However, the neuroscientific foundations of AP are still far from being understood. It is currently generally assumed that a network comprising many brain regions underlies AP.

One of the most important perisylvian brain regions for AP is the planum temporale (PT), which is involved in auditory processing, especially in categorical perception (Griffiths & Warren, 2002; Schlaug et al., 1995). A previous structural study revealed that AP musicians have a smaller right-sided PT surface (Keenan et al., 2001), which leads to a stronger left surface asymmetry in APP (Griffiths & Warren, 2002; Schlaug et al., 1995). Moreover, previous functional studies reported positive correlations between pitch-labeling ability and left-sided PT activation (Wilson et al., 2009; Zatorre et al., 1998), while the right-sided PT has been suggested to be mostly involved in pitch perception (Burkhard et al., 2019; Loui et al., 2011; Wengenroth et al., 2014). Heschl's gyrus (HG), the superior temporal sulcus (STS), and the middle temporal gyrus (MTG) are other perisylvian brain regions that have also been associated with AP perception and categorization (Burkhard et al., 2019; Loui et al., 2011; Wengenroth et al., 2014). Furthermore, a recently published study revealed enhanced intracortical myelination in the right planum polare for APPs compared to non-APPs (Kim & Knösche, 2016). In summary, AP seems to be modulated by many different perisylvian brain areas, which form, at least partially, an interconnected network. Compatible with this view, several functional and anatomical studies have shown hyperconnectivity in perisylvian brain regions for APPs (Jäncke et al., 2012; Kim & Knösche, 2017; Loui et al., 2011, 2012).

Another important brain structure is the dorsolateral frontal cortex (DLFC). This structure is assumed to be involved in the pitch-labeling process, enabling associative learning in APP (Bermudez & Zatorre, 2005; Schulze et al., 2009). Moreover, it has been suggested that AP ability is subserved by different brain regions involved in working memory: for example, the posterior inferior frontal gyrus, premotor areas, as well as the inferior parietal lobe (Schulze et al., 2009).

1.4 Open questions

"To understand how biological systems work, we must make an observation, ask why it came about, form a hypothesis, design and perform an experiment that will either support or refute that hypothesis, and, finally, draw a conclusion. (...) ... it is the only way that a topic can move along on sure footing." (Gazzaniga, 2013, p. 6)

As described in the preceding chapters, previous studies have provided many interesting insights regarding AP ability. Nonetheless, some uncertainties remain regarding the psychological and neuroscientific foundations. An important prerequisite to robustly detect even small true effects is to ensure high statistical power. Furthermore, the higher the power, the smaller is the probability of detecting an effect when there is no true effect in the investigated population. Therefore, high power increases the confidence in the results of a given study. Power increases with increasing sample size. Since AP is a relatively rare phenomenon, most of the previous studies inevitably worked with relatively small sample sizes (usually around 20 participants per group). For this reason, studies with higher sample sizes are essential to reveal the neural underpinnings of AP in the long run. With this fundamental principle in mind, this doctoral thesis examines AP within a relatively large sample, including more than 100 participants with different degrees of AP.

An important question for AP research is whether the characteristics of the ability derive from perceptual or rather cognitive processes. Furthermore, it is also possible

that AP ability is only enabled through a specific interaction of perceptual and cognitive processes. There are different opinions on this issue in the existing literature (Hirata et al., 1999; Levitin & Rogers, 2005; Levitin, 1994; Schulze et al., 2009; Wengenroth et al., 2014)**.** The question remains of how it is possible to distinguish between these two processes and how these processes can be easily and reliably operationalized according to Ockham's razor. Thanks to basic research, we know that perceptual processes occur at a relatively early stage of neuronal processing and that cognitive steps are built upon them and therefore necessarily occur at a later stage (Koelsch & Siebel, 2005). Moreover, we know approximately which brain areas are expected to be involved in the auditory processing chain. For example, we know from previous studies that auditory processing around 100 ms after stimulus onset occurs in the primary auditory cortex, also referred to as HG and that activation of more cognitive processes are distributed over several brain areas mostly located in the frontal and parietal lobes (Koelsch, 2005; Näätänen & Picton, 1987; Verkindt et al., 1995).

However, the studies presented here focus in particular on which processing stages (early or late) and which related brain areas are pivotal for AP processing. To answer these questions, two experiments were conducted using either the functional electroencephalography (EEG) technique (Study I) or a structural diffusion tensor imaging (DTI) paradigm (Study II). The EEG technique applied in Study I is a good method to investigate the processes underlying the temporal aspects of AP ability because it combines an excellent temporal resolution with reasonable spatial resolution. In Study II, we were interested in white matter (WM) features underlying the brain areas that are considered as important for AP ability. Therefore, we applied a DTI design.

Study I: The primary objective of Study I was to identify the time frames during which AP processing differs from RP processing while listening to sine tones and noise segments. The intention was to keep the listening conditions as ecologically valid as possible. Therefore, an attentive listening paradigm was applied. However, to be able to

answer this question, a method providing excellent temporal resolution was required. The ideal method was therefore EEG, or more precisely, the investigation of different event-related potentials (ERPs). Here, a distinction is made between early auditory evoked potentials, primarily reflecting perceptual and sensory processes, and later ERPs, which are mostly associated with cognitive processing stages (Baumann et al., 2008; Elmer et al., 2013; Kutas & Federmeier, 2011; Shahin et al., 2003, 2005; Tremblay et al., 2001; Vaughan & Ritter, 1970).

Study II: Study II exclusively addresses the structural investigation of the WM underlying perisylvian brain areas. We examined the question of whether fractional anisotropy (FA) values of different brain areas vary as a function of AP proficiency level. In order to answer this question, we applied a correlational approach in association with DTI. With this approach, we avoided an arbitrarily determined group classification to account for the continuous manifestation of AP ability (Bermudez & Zatorre, 2009; Oakes, 1955).

1.5 Methods

At present, many excellent techniques exist for the in vivo investigation of the human brain. As is often the case, each methodical approach has advantages and disadvantages. Therefore, the choice of the method used depends on the question that is asked.

In Study I, we addressed a question that could only be addressed with high temporal resolution. Hence, EEG was the first choice, as it has a temporal resolution in the millisecond range. The most important drawback of EEG is its relatively moderate spatial resolution. However, mathematical solutions enabling the computation of the neuronal generation of the EEG provide a remarkably precise estimation of the underlying cortical sources. Additionally, the spatial resolution can be improved by using more electrodes, for example, provided by so-called high-density EEGs (Jäncke, 2005).

However, if the question raised demands an answer couched in spatial terms, another technique must be applied. As already mentioned, the aim of Study II was to identify WM structures underlying different brain areas that relate to AP ability. To investigate this issue, DTI is clearly the method of choice.

In the following sections, the methods used in the studies presented in this doctoral thesis are described. As the method used in Study I (the so-called signed area) is not yet fully established and therefore less frequently used in previous studies, a more detailed explanation of this method is provided. Otherwise, the method used in Study II (tract-based spatial statistics; TBSS) is relatively popular and widely accepted and is therefore described only in general terms.

1.5.2 The origin of the EEG signal and source reconstruction

Unless otherwise stated, the following section refers to Luck (2014). Electrical activity can be produced in two different ways by a neuron: either due to an action potential or by postsynaptic potentials (PSPs). Mostly, ERPs reflect PSPs and not action potentials (except some very early auditory brainstem potentials). The reason is that for detection at the scalp level, a summed neuronal activity is necessary. However, due to the positions of the neuronal axons and the timing of the action potentials, it is almost impossible to record activity derived from action potentials within a few milliseconds of stimulus-onset. This is, because the signals from individual action potentials will cancel each other out if the neurons do not fire at exactly the same time, which is usually not the case in large neuronal populations. In contrast, PSPs originate from a graded change in voltage caused by ion channels that open and close in response to neurotransmitters binding to the receptors of the postsynaptic cell. The signal detected by noninvasive scalp electrodes, using the EEG technique, is primarily generated by those PSPs. Usually, a complete electrode montage contains 20 to 256 electrodes that pick up electrical signals from the cortex and to a lesser extent from the subcortex. The captured activity is then amplified and recorded as voltage changes over time. A simplified

representation of an EEG recording and coarse preprocessing steps are depicted in Figure 1.1.

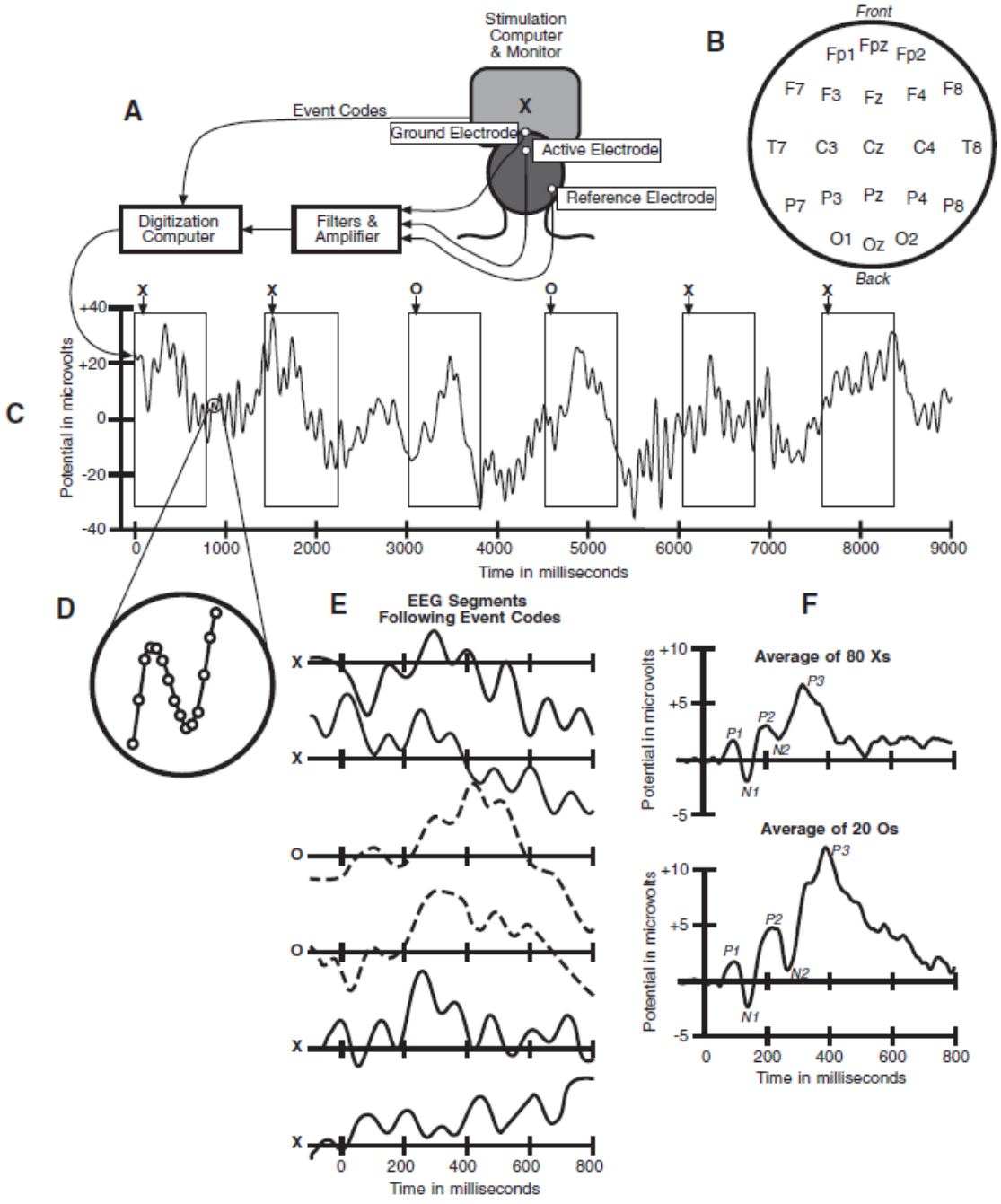

Figure 1.1: Broad overview of an ERP paradigm. A standard montage according to the International 10/20 System includes several active, ground, and reference electrodes (A,

B). After filtering and amplifying, the signal is converted to distinct digital samples (C, D). To precisely track the stimulus onsets in time, event markers enable the identification of each stimulus presentation (E). After segmentation, single trials of the same stimulus are averaged to get activation time locked to the stimulus, in other words: event related (F). (Luck, 2014; used with friendly permission from MIT press).

One of the main problems concerning the EEG technique is the estimation of the sources that generated the EEG signal at the scalp level. The so-called forward problem describes the fact that if the sources of the EEG signal were known, one could calculate the exact potential distribution on the scalp level. In the context of EEG, we are presented with the opposite pattern: the potential distribution on the scalp level is perfectly known but we would like to estimate the underlying generators – this is known as the inverse problem. For the inverse problem, there is no straightforward solution as there is for the forward problem, but the problem has an infinite variety of solutions.

1.5.1 Event-related potentials

Although the analysis of the EEG signal gives the advantage of a very high temporal resolution, the signal cannot be used in its raw form. The EEG signal is a very broad measure of brain activity and is comprised of an accumulation of different signals derived from different sources. To be able to extract neuronal processes that are related to specific events a simple averaging technique has to be applied. In particular, this means that all EEG segments of the same kind of trail are aligned relative to stimulus-onset and averaged. The more segments enter the averaging procedure, the better the signal-to-noise ratio. The idea behind the averaging process is simple: it is based on the assumption that the cortical process that is triggered by a specific stimulus stays more or less the same if the same stimulus is repeatedly presented. In contrast, brain activity that is randomly aligned with rather than time-locked to the stimulus is canceled out (Jäncke, 2005). This simple process allows the isolation of electrical potentials that are associated with specific events, the so-called ERPs (Fig. 1.1F). Although ERPs do not

provide us with direct information about the human mind, they enable us to answer temporal questions with a precision that cannot be achieved with any other neuroscientific method.

An ERP component has either a positive or negative polarity and is usually named after the time it appears after stimulus onset or after their order of appearance. Every component has a quantifiable latency and amplitude. As a rule, the number of active neurons is positively related to the amplitude (Jäncke, 2005). It is important to keep in mind that, in general, the amplitude is operationalized as voltage relative to the average voltage in the pre-stimulus time period (baseline). Since in Study I the amplitudes of the ERP components were investigated, the next paragraph explains the different analysis possibilities in more detail.

This paragraph refers to Luck (2014) unless specified otherwise. A primary advantage of analyzing the amplitude of an ERP component is that it is a simple and robust approach in comparison to other more sophisticated analysis methods. One of the most important and also most challenging matters is the choice of a time window during which the amplitude is quantified. The definition of a measurement window can bias the statistical analysis toward significance: for example, when choosing a time window with the largest pronounced differences (between groups, conditions, etc.). In the following, some of the most popular methods for analyzing the amplitude of an ERP component are listed and described.

Peak amplitude

In the past, when computers were more expensive or did not yet exist, measuring the so-called peak amplitude was the method of choice for analysing EEG data. The procedure of this method is as follows: after defining a suitable time window in which to measure the component of interest, the maximum voltage point within this time window is located. This is called the simple peak amplitude. As shown in Figure 1.2A, taking the maximal voltage in a predefined time window does not always lead to an appropriate

measure of the component of interest. If a time window of 150–300 ms is chosen for finding the maximum voltage to operationalize the amplitude of the P2 component, the maximum point lies on the border and represents the onset of the P3 component rather than the P2 peak amplitude. Defining a narrower time window is not a potential solution because of the peak latency variation (e.g. across subjects). A slightly modified version of the simple peak amplitude is the local peak amplitude. This method is preferable because it requires the maximum voltage to be surrounded by lower voltages on both sides. As depicted in Figure 1.2A, the local peak amplitude is a better representation of the P2 component. Many software packages – for example, the peak detection implemented in BrainVision Analyzer – compute the local peak amplitude.

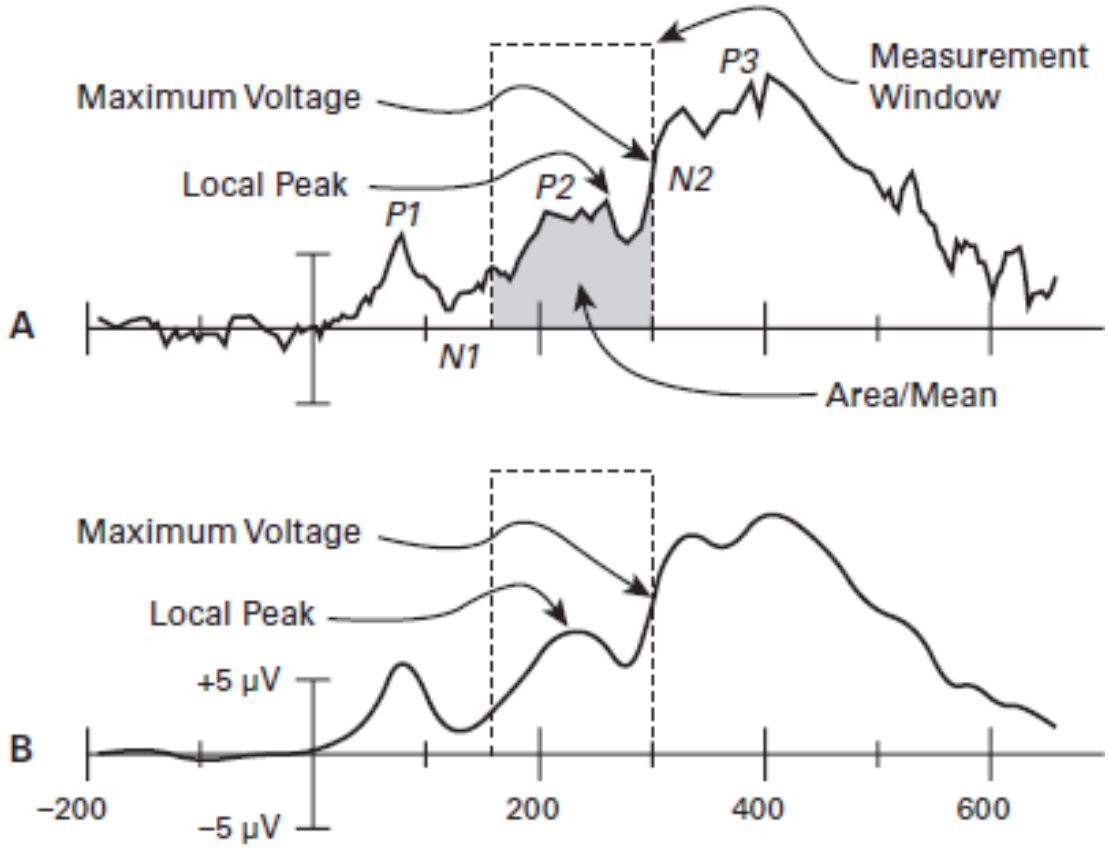

Figure 1.2: Measuring the P2 wave within a predefined time window of 150–300 ms after stimulus onset. (A) Illustration of the maximum voltage at the borders of the predefined time window (300 ms) and the local peak. Furthermore, the region displayed

in gray can be used to calculate area or mean amplitude. (B) To attenuate high-frequency noise, a high-cutoff filter is applied to the signal (Luck, 2014; used with friendly permission from MIT press).

Mean amplitude

As the name implies, the mean amplitude is calculated by defining a time window and computing the average voltage over all sample points during this window.

Area amplitude

The geometric term area can describe a couple of things and therefore has to be defined. Using the examples depicted in Figure 1.3, the usage of the geometric term area in ERP analysis is explained. First of all, the unit of area is a multiplication of the two involved dimensions: voltage and time. Hence, for ERP analysis the unit is given in μV * ms. As shown in Figure 1.3A, both regions (A and B) are bounded by the edges of the defined time window – in this case, 200 to 300 ms – and the baseline. An important aspect is that area is always positive and therefore will sum (see the rectified waveform in Fig. 1.3B). In contrast to the mean amplitude, the positive and negative deflections do not cancel each other out. Area is not identical to the integral, in which one region is subtracted from the other (B–A). If the integral is divided by the duration of the time window (in this case 100 ms), we obtain the mean amplitude. Usually, neither the rectified area nor the integral is used in ERP research. In ERP research, we are interested in either measuring a positive or a negative component. Hence, we should either use the area above or below the baseline, respectively. This is exactly what is done if ERP components are extracted by using signed area (Fig. 1.3C).

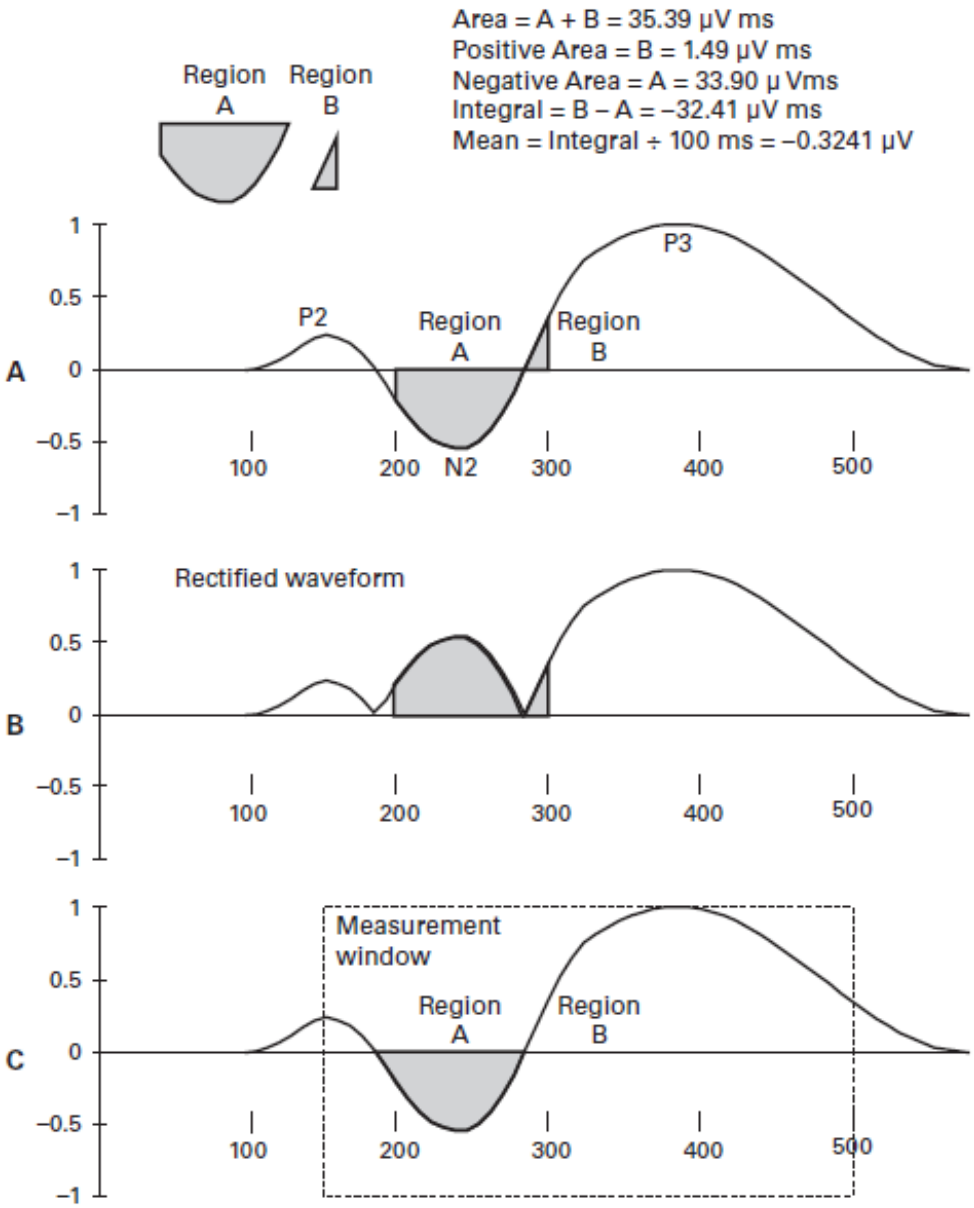

Figure 1.3: Definition and calculation of area amplitude. (A) A time window of 200–300 ms is used to define the two regions A and B. As a geometric shape, bounded by the baseline and the ERP waveform itself, the term area is always positive (µV * ms). Panel (B) shows the same waveform as panel (A) but in a rectified form, which means that negative voltages were converted into positive voltages by multiplying them with -1. (C) One of the main advantages of calculating area amplitude is that a wide measurement window can be chosen because regions of undesired polarity are excluded (Luck, 2014; used with friendly permission from MIT press).

Although brain plasticity can be tracked in many different ways (e.g. cortical thickness or acetylcholine levels), a promising approach is the investigation of WM properties such as axonal connectivity. WM has an important role in information transmission between different brain regions. Furthermore, changes in myelination could have an important influence on the timing of information transmission (Bennett, 1971; Dan & Poo, 2004; Markham & Greenough, 2004), which in turn influences cognitive functioning. A promising approach to draw conclusions about the WM structure of the human brain is the DTI technique developed in the mid-1990s (Basser et al., 1994). DTI is able to visualize and quantify diffusion within the brain and neural tissue in general. Importantly, DTI measures represent an indirect marker of biological tissue because they rely on the thermal (Brownian) motion of water molecules within the tissue itself. If water molecules diffuse in a uniform fashion such that they disperse in an equidistant manner in all directions, the process is called isotropic diffusion. In contrast, if obstructions hinder the diffusion of the water molecules in an isotropic manner and instead force them to diffuse along one axis, the diffusion is termed anisotropic. The direction and degree of diffusion are measured to draw conclusions about the environment of the diffusing water molecules, or simply, to infer the corresponding biological structure. A common way to assess brain connectivity is so-called tractography, which is not part of the present thesis. In short, the principal direction of water diffusion is estimated, which in turn is assumed to correspond to the underlying orientation of axons. Briefly, the diffusion coefficient depends on the expected signal loss, which in turn is the result of translational motion of water molecules. When exposed to the homogeneous MRI magnetic field, all water molecules are aligned and yield a signal of equal frequency. Introducing a gradient pulse, along any desired orientation, linearly alters the strength of the MRI magnetic field and the water molecules drift out of phase, resonating at different frequencies along the gradient axis. Afterward, the signal from the water molecules is out of phase (depending on the location) but again begins to resonate at the same frequency as before the gradient was

applied. Introducing a second gradient, typically 20–50 ms after the first gradient, of opposite polarity unwinds or refocuses the phase. As shown in Figure 1.4, due to the motion of water molecules, the refocusing is not perfect but rather leads to a signal loss (depending on the timing and strength of the gradient pulse) and a higher diffusion coefficient.

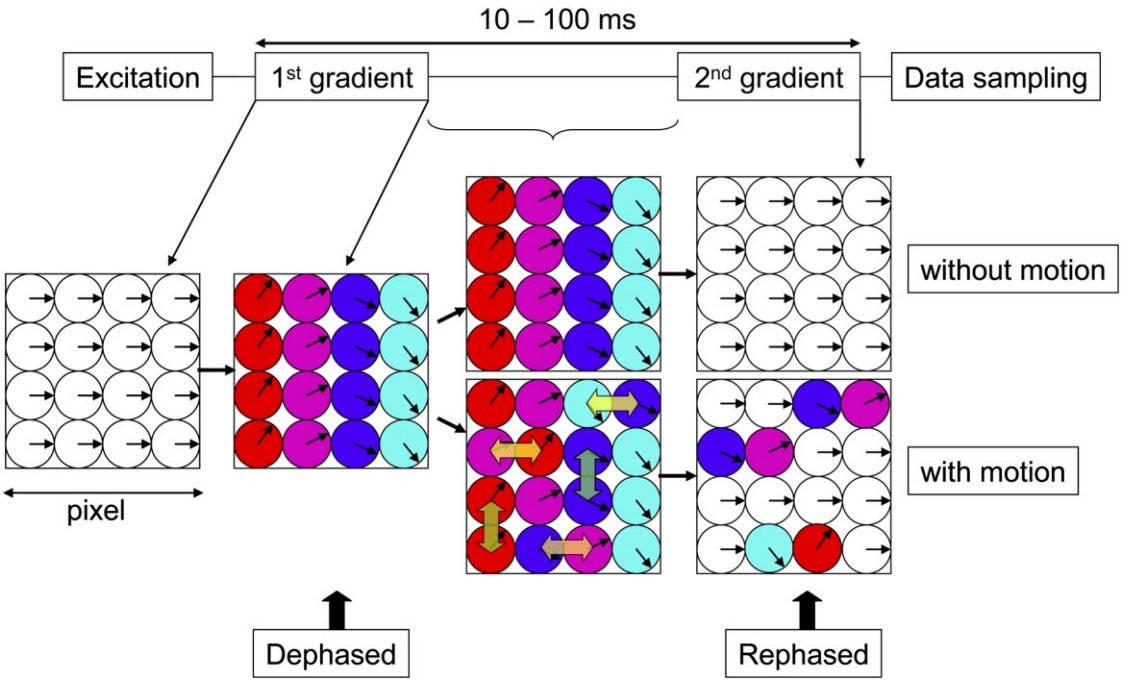

Figure 1.4: Motion of water molecules after gradients are applied. The circles depict different water molecules with the vectors representing the particular phases. Those phases are not able to be perfectly re-focused by applying the second gradient if the water molecules have moved between the first and the second gradient application. This leads to signal loss (Mori & Zhang, 2006; used with friendly permission from Elsevier).

Furthermore, an apparent diffusion coefficient map can be created by calculating the diffusion coefficient for each voxel separately. Since the orientation of fibers is mostly slanted with respect to the X, Y, and Z axes, accurate estimation of the largest apparent diffusion coefficient orientation would require measuring the diffusion along thousands of axes. Clearly, this is impractical, and a model was therefore introduced to fit measurements along the different axes to a 3D ellipsoid, which is called the diffusion

tensor (Basser et al., 1994). At each voxel, six parameters were estimated, which can be used to generate different contrasts (Fig. 1.5).

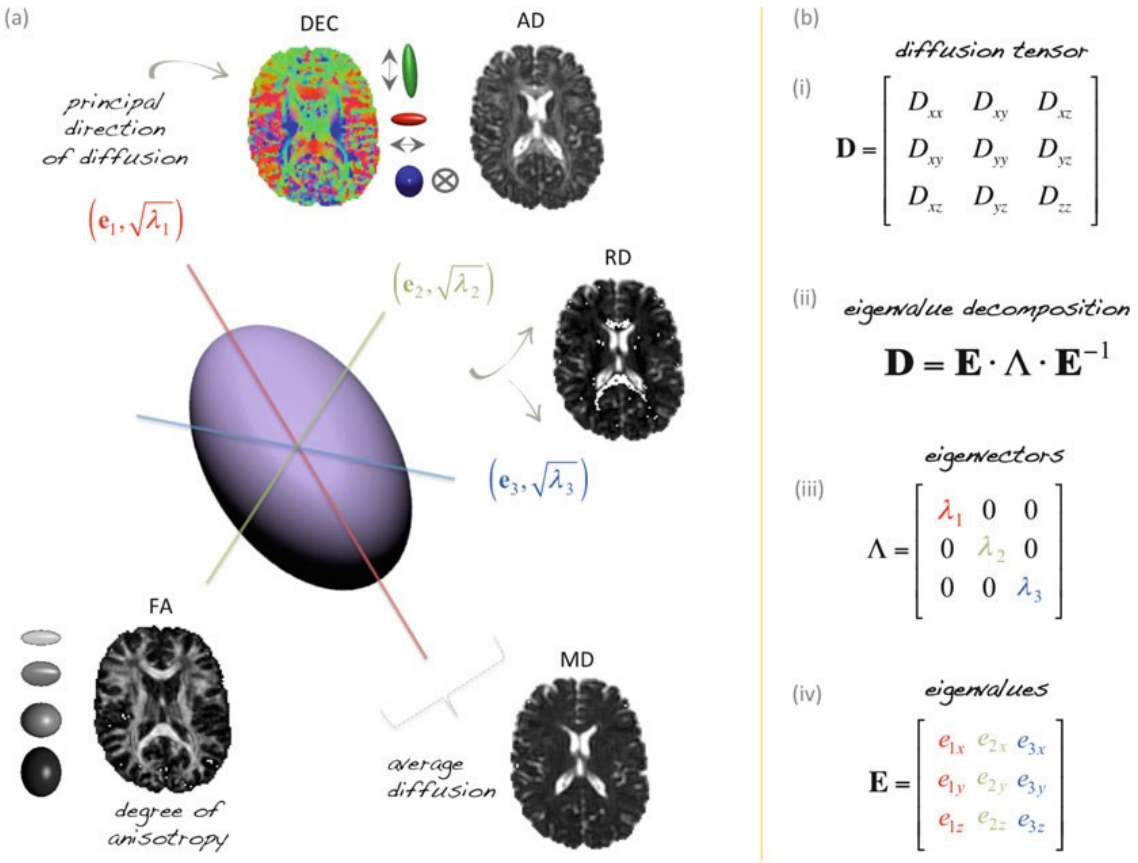

Figure 1.5: The concept of a diffusion tensor. The average diffusion distance in all directions is depicted by the ellipsoid. Mathematical decomposition of the tensor: whereas the three eigenvectors describe the direction, the three eigenvalues represent the intensity of diffusion. The first eigenvector describes the so-called axial diffusivity, which is simply the direction of the maximal diffusion. The radial diffusivity value is obtained via the second and third eigenvalues. The average of all eigenvalues is called mean diffusivity, which makes no assertion about the direction of diffusivity. Finally, FA represents the degree of anisotropy, ranging from 0 (isotropic) to 1 (anisotropic) (Van Hecke et al., 2015; used with friendly permission from Springer Nature, p. 12).

However, a widely used metric to investigate the degree of anisotropy in WM is the FA, which is thought to increase with myelination (Fields, 2005). Although FA is sometimes interpreted as WM integrity or even as an indicator of the WM fiber tract quality (Alexander et al., 2007), it is important to mention that its computation is based on an oversimplified model, which relies on the diffusion of water molecules averaged over a whole voxel. Therefore, an appropriate interpretation considers FA as an indirect measure of microstructural tissue status (Alexander et al., 2007; Van Hecke et al., 2015). A common approach to investigate FA presented below, is TBSS, which operates on a voxelwise basis to calculate either differences across groups or correlational analysis.

1.6.1 TRACT-BASED SPATIAL STATISTICS

According to Smith (2006), FA is a good measure when the aim is a comparison of subjects at the voxel level. The reasons for this are, inter alia, that FA is a relatively objective measure that is mostly independent of the local WM fiber structure. The TBSS application generally aims for easier interpretation and greater objectivity and sensitivity of the DTI analysis. Furthermore, the aim of the TBSS method is to overcome the problems that other methods (such as voxel-based morphometry or tractography) show if applied to DTI data, such as alignment inaccuracies, the arbitrary choice of smoothing extents, and the need for hand-drawn ROIs. TBSS allows for whole-brain investigation, which does not necessarily require the specification of predefined ROIs. Moreover, TBSS requests a minimum of user intervention, making the application more automatic and therefore less user-dependent. In short, the preprocessing comprises the pre-alignment (correction for head motion and eddy currents), the estimation of the diffusion tensor, and the exclusion of non-brain voxels from further analysis. In order to maintain the fundamental topological structure and to enable a local comparison, the aligning process (voxelwise non-linear registration with intermediate degrees of freedom) is driven by the FA images itself. The next objective is to project the data onto the so-called tract skeleton. To apply the alignment, the most

typical subject is selected by calculating the amount of warping needed to achieve a minimum mean distance to all other subjects of the investigated sample. After the FA images of all other subjects are aligned to the selected image, an affine transformation into the MNI152 space (1 x 1 x 1 mm^3) is carried out for reasons of comparability. A mean FA image is further created and used to generate the tract skeleton. This skeleton represents the center of all tracts that are common to all subjects in the sample, with the FA values becoming gradually smaller when moving away from the tract center. To exclude non-WM voxels from further analysis, the FA skeleton is thresholded (usually between a mean FA of 0.2–0.3). As a result, a good tract correspondence across the different subjects in the sample is achieved. In the next step, the individual FA images are projected onto the mean FA skeleton, assigning the individual maximum FA values to the skeleton voxel. As the last step, voxelwise statistics across subjects are applied for each voxel on the skeleton. To correct for multiple comparisons, permutation-based methods are usually applied.

2. EMPIRICAL PART STUDY I

Early tone categorization in absolute pitch musicians is subserved by the right-sided perisylvian brain

Anja Burkhard[1*], Stefan Elmer[1], Lutz Jäncke[1,2]

[1] Division Neuropsychology, Institute of Psychology, University of Zurich, 8051 Zurich, Switzerland
[2] University Research Priority Program (URPP), Dynamics of Healthy Aging, University of Zurich, 8051 Zurich, Switzerland

This article was originally published in:

Burkhard, A., Elmer, S., & Jäncke, L. (2019). Early tone categorization in absolute pitch musicians is subserved by the right-sided perisylvian brain. *Scientific Reports, 9(1)*, 1–14.
doi: 10.1038/s41598-018-38273-0

2.1 Abstract

Absolute pitch (AP) is defined as the ability to identify and label tones without reference to keyality. In this context, the main question is whether early or late processing stages are responsible for this ability. We investigated the electrophysiological responses to tones in AP and relative pitch (RP) possessors while participants listened attentively to sine tones. Since event-related potentials are particularly suited for tracking tone encoding (N100 and P200), categorization (N200), and mnemonic functions (N400), we hypothesized that differences in early pitch processing stages would be reflected by increased N100 and P200-related areas in AP musicians. Otherwise, differences in later cognitive stages of tone processing should be mirrored by increased N200 and/or N400 areas in AP musicians. AP possessors exhibited larger N100 areas and a tendency towards enhanced P200 areas. Furthermore, the sources of these components were estimated and statistically compared between the two groups for a set of a priori defined regions of interest. AP musicians demonstrated increased N100-related current densities in the right superior temporal sulcus, middle temporal gyrus, and Heschl's gyrus. Results are interpreted as indicating that early between-group differences in right-sided perisylvian brain regions might reflect auditory tone categorization rather than labelling mechanisms.

2.2 Introduction

Absolute pitch (AP) musicians are able to identify tones without relying on an external or internal reference tone. AP is a rare ability for which some authors estimate the occurrence of 1 in 10.000 people (Takeuchi & Hulse, 1993). However, recent large-scale studies have shown that the incidence of AP ability must be much higher particularly in professional musicians and is influenced by several factors (e.g., age of commencement, ethnicity, tonal language background, and anatomical prerequisites (Deutsch et al., 2006; Deutsch et al., 2009)). The neural and cognitive underpinnings of AP are still a matter of debate. For example, it remains unclear whether this ability

originates from hereditary transmission, early exposure to environmental factors, or both (for a summary (Zatorre, 2003)). Another debate is concentrated around the question of which processes account for this very specific ability. An absolute pitch possessor (APP) is able to categorize, label, and produce a specific tone without making use of a reference tone (Bachem, 1955; Miyazaki, 1989; Takeuchi & Hulse, 1993), while a relative pitch possessor (RPP) relies on the latter. AP is especially useful in categorizing the chroma of a given tone and not necessarily its height (Bachem, 1955; Miyazaki, 1989). Thus, APPs are thought to experience pitches more categorically than RPPs or non-musicians. However, it is still a matter of debate whether AP describes an all-or-nothing or rather a gradually distributed ability. Already Bachem (1937) introduced the term quasi-AP, which describes the ability to estimate a given tone based on a single internal reference tone that is available. This internal reference might derive either from the deepest tone that a musician is able to sing or, for example, from the tuning tone for a given instrument. Based on such an internal reference, tone intervals can be calculated and the target tone can be estimated accordingly. Furthermore, Wengenroth et al. (2014) reported that a partial AP ability in musicians is more frequent than a perfect AP or no AP ability. This leads to the difficulty to fully disentangle the influences of these two abilities, possibly resulting in an overestimation of the AP ability. However, a consistent question in AP research is whether early perceptual stages of tone processing account for this ability, or whether later stages are the important ones (or both). The early perceptual encoding stages were suggested to be important by several previous studies (Hirata et al., 1999; Schulze et al., 2009; Wengenroth et al., 2014). In particular, Schulze et al. (2009) proposed that the difference between APPs and non-APPs lies in a specific perceptual ability of AP musicians that relies on the encoding of tonal information according to predefined pitch chroma categories. Otherwise, the labelling process is assumed to be a second component that is based on learned associations. In this context, Levitin (1994) proposed a two-component AP model where the ability of AP can be explained by pitch-memory and pitch-labelling. Pitch-memory is thought to be widespread even in non-APPs, whereas pitch-labelling is

considered to be the main component accounting for AP ability (Levitin & Rogers, 2005; Levitin, 1994).

Due to the high temporal resolution of EEG or MEG, auditory evoked potentials represent an ideal biomarker for examining both early and late processes. The so-called early auditory evoked potentials (N100, and P200) primarily reflect sensory and perceptual processes at the initial stages of auditory analysis (Hillyard et al., 1998) and are thought to be generated by the auditory cortex (Elmer et al., 2013; Vaughan & Ritter, 1970). Furthermore, several studies were able to associate the N100/P200 complex with improved perception (Baumann et al., 2008; Seppänen et al., 2012; Shahin et al., 2003, 2005; Tremblay et al., 2001). By contrast, later components occurring between 200–1,500 ms after stimulus onset have rather been associated with endogenous and cognitive processes (Kutas & Federmeier, 2011). In this vein, auditory categorization (Dittinger et al., 2016), contextual influences, and lexical selection (Kutas & Hillyard, 1980; van den Brink et al., 2001) have been associated with N200 responses, while the N400 component can be used as a marker for semantic categorization (Kutas & Hillyard, 1980), episodic memory functions (Besson & Kutas, 1993) as well as pitch labelling processes (Elmer et al., 2013).

In a previous electrophysiological study conducted with AP and non-AP musicians, Wu et al. (2008) examined the N100 component under three different listening conditions. In one condition, participants had to listen passively to sine tones, in a second condition they made relative pitch (RP) judgments, and in a third condition, they had to label tones without using a reference. Even though the authors revealed global field power differences between the different conditions during the time window of the N100 component, they did not observe between-group differences. However, the intracortical sources of the N100 component were spatially more extended in the left and right auditory regions in the AP group during the labelling condition. Other studies that focused on the N100 component and their intracortical sources (Hirose et al., 2004) reported bilateral increased N100 responses in the auditory cortices among APPs during

a labelling task, whereas the N100 component only increased in the left hemisphere for RPPs. However, it is still a matter of debate whether auditory cortex activity is differentially modulated within the AP group as a function of attention. Hirose and colleagues (2004) postulated that the right hemisphere is involved in pitch height analysis, while the left hemisphere is involved in label assignment. Furthermore, Pantev et al. (1998) found no differences in the N100 component between AP and non-AP musicians during a passive listening task, even though only the left hemisphere was examined. However, when they compared musicians against non-musicians, they observed stronger dipole moments for piano tones in the musician group. Furthermore, Itoh et al. (2005) suggested the existence of a so-called AP negativity, which occurred around 150 ms after stimulus onset in high-scoring APPs only. This AP negativity was observed during listening, pitch-naming, and auditory Stroop tasks. Furthermore, the amplitude of the AP negativity was modulated by stimulus congruence in the auditory Stroop task. Otherwise, differences in the P200 component have been observed during auditory feedback (Behroozmand et al., 2014) as well as during passive listening (Wengenroth et al., 2014) tasks with APPs showing larger amplitudes. Finally, Elmer and colleagues (Elmer et al., 2013) found no group differences between APPs and RPPs during a passive listening task for the N100 and P200 components. However, the authors reported increased N400 amplitudes for APPs during pitch-label association tasks.

While the aforementioned studies used EEG or MEG, other studies examined the functional anatomy of AP by means of fMRI and PET. When listening to tones compared to noise bursts, APPs revealed stronger blood flow in the left posterior dorsolateral frontal cortex (Zatorre et al., 1998), leading the authors to assume that this region is involved in verbal-tonal association learning. Furthermore, when comparing an interval judgment task with noise bursts, APPs demonstrated stronger blood flow in the right middle and inferior temporal cortex, which are regions suggested to facilitate multi-modal processing (Zatorre et al., 1998) and auditory object recognition (Rauschecker, 2015). The left dorsolateral prefrontal cortex was also associated with AP

ability in a study by Ohnishi et al. (2001), who reported correlations between AP performance and hemodynamic responses in the left dorsolateral prefrontal cortex and planum temporale. Furthermore, stronger hemodynamic responses in the middle part of the left superior temporal sulcus (STS) were also observed in APPs during a pitch memory task (Schulze et al., 2009). In contrast, RPPs demonstrated stronger hemodynamic responses in the right superior parietal lobe and intraparietal sulcus. In general, the right auditory areas are thought to be more strongly involved in spectral analysis (Hirose et al., 2004; Zatorre & Belin, 2001), while the left auditory areas are suggested to be more strongly involved in rapid temporal processing (Zatorre & Belin, 2001) and the assignment of tone labels (Hirose et al., 2004).

In this study, we examined the effects of attentive tone listening on classical ERPs, which have often been used to study auditory functions and related cognitive functions. In a second step, we explored the ERP-related intracortical sources and statistically compared them between the two groups by selecting a priori defined regions of interest (ROIs). This approach is particularly fruitful, especially when the data are collected using a high-density EEG array of 128 channels. Since we were interested in tracking the neural underpinnings underlying AP ability and not in differences between musicians and non-musicians, we compared musicians with AP to musicians without AP. A further innovative aspect of our study is that we measured a relatively large sample of 103 participants, whereas prior AP studies often used small sample sizes which makes it difficult to detect a true effect and increases the likelihood of reporting false-positive findings. Furthermore, since the labelling process is suggested to occur automatically in APPs (Levitin & Rogers, 2005) an attentive listening task was used instead of a labelling task. In this context, participants were not forced to explicitly label tones, enabling a listening experience that is as natural as possible. Moreover, we used sine tones that were either in tune or slightly mistuned in order to prevent supportive instrumental cues.

In this EEG experiment, we addressed two different research questions. Firstly, we wanted to determine the time windows reflecting between-group differences, as this reveals the rough nature of the underlying perceptual and cognitive processes. In fact, early differences in the N100 and P200 components indicate an AP-related perceptual tone processing mode at the level of the auditory cortex. Based on previous studies, we expected higher N100/P200 amplitudes for the AP group compared to the RP group (Behroozmand et al., 2014; Hirata et al., 1999; Wengenroth et al., 2014; W. Wu et al., 2008). Furthermore, we assumed that putative between-group differences in the N200 and N400 components would reflect a differential engagement of cognitive processes such as lexical selection, memory functions, and tone-label associations (see for example Elmer et al., 2013). Secondly, we expected that only APPs would be aware of the slightly mistuned tones since it has often been reported that APPs react disturbed and annoyed when recognizing mistuned acoustic information (Levitin & Rogers, 2005). Therefore, we expected differential neuronal responses for APPs while listening to mistuned tones.

2.3 Materials and Methods

2.3.1 Subjects

In the present study, we measured 103 musicians with and without AP. Nine participants had to be excluded due to drug abuse and/or psychiatric disorders (depression and anxiety). All participants were tested with pure-tone audiometry (MAICO Diagnostic GmBh, Berlin) in the frequency range of 250–8000 Hz. According to this procedure, all participants demonstrated a normal and non-pathological audiological status (i.e., all tested frequencies could be heard below a threshold of 30 dB). All participants gave informed consent and were paid for participation. The study was approved by the Ethics Commission of the University of Zurich in accordance with the declaration of Helsinki. All participants were highly trained musicians with a comparable proficiency level to avoid confounding factors regarding

musical experience. To prevent an arbitrary classification criterion based on the AP-test score (see below) participants were divided into two groups based on self-report. The 49 participants who claimed to possess AP were grouped into the AP group and the remaining 45 musicians were assigned to the RP group. In the AP group, we examined eight wind players, 29 string players, one percussionist, nine pianists, and two singers. In the RP group, we included twelve wind players, 19 string players, five percussionists, seven pianists, and two singers. As depicted in Table 1, groups were comparable in gender, handedness (Annett, 1970; Bryden, 1977), the occurrence of synaesthesia, and bilingualism. Furthermore, musical aptitude was evaluated using the AMMA test (Gordon, 1989). Applying an additional in-house test, participants had to name ten intervals of two sequential sine tones of different frequencies to evaluate RP ability. This short procedure was applied as a screening test to confirm that all participants were characterized by RP abilities. General cognitive ability was evaluated using a standard German IQ screening test (KAI: Kurztest für allgemeine Intelligenz (Lehrl et al., 1992)). The age at which participants started to play an instrument, as well as the estimated training hours accumulated over the lifespan, were surveyed (Table 2).

Table 2.1: Gender (masculine/feminine), handedness (right/left/both), occurrence of synaesthesia (yes/no) and bilingualism (yes/no) proportions for both groups (AP/RP) are displayed.

	Gender	Handedness	Synaesthesia	Bilingualism
	(m/f)	(r/l/b)	(y/n)	(y/n)
AP N = 49	26/23	42/4/3	13/36	10/39
RP N = 45	20/25	40/4/1	13/32	15/30

Table 2.2: Means, standard deviations, and p-values (t-test, two-tailed) depicted for age, musical aptitude, correctly identified intervals, cognitive capability, age of commencement of instrumental training, and cumulative number of training hours across the lifespan.

	Age	AMMA	Interval	KAI	Age of Commencement	Training Hours
AP						
N = 49	27.43 (4.85)	66.37 (6.05)	7.71 (1.59)	122.74 (33.24)	6.02 (2.34)	15‘445 (12‘248)
RP						
N = 45	26.16 (4.58)	64.11 (6.66)	7.60 (2.03)	130.65 (29.47)	6.49 (2.51)	13‘359 (9‘705)
p-value	.195	.088	.761	.230	.352	.365

2.3.2 AP TEST

In order to verify AP ability, pitch labelling accuracy was also tested by a modified version of a previously used in-house developed test (Oechslin et al., 2010) that each participant completed at home. With this procedure, we could assure that participants who claimed to possess AP had a corresponding score. The test consisted of 108 pseudo-randomly presented pure sine wave tones ranging from C3 to B5 (A4 = 440 Hz). After the presentation of each tone, the participants had to click on the appropriate label presented on a screen. Each tone trial had a time limit of 15 seconds. Immediately before and after each tone, two seconds of Brownian noise was presented to prevent mnemonic cues. If the participants did not select the corresponding tones in the time window of 15 seconds, the test continued with the next trial. After participants had accomplished the AP test they received an invitation for the EEG measurement. AP scores were obtained by computing the percentage of correct answers in the AP test (Fig. 1). An answer was classified as correct if the exact pitch chroma was identified. Octave errors were neglected because APPs are usually able to identify the chroma but not necessarily the height of a tone (Bachem, 1955; Miyazaki, 1989; Takeuchi & Hulse, 1993). Afterward, a t-test (two-sided) was computed to uncover group differences. AP test performance substantially differed (as expected) between both groups ($p < .001$, Cohen's $d = 2.66$, mean ± sd for the AP group: 76.43 ± 20.15, RP group: 24.91 ± 18.53). The performance of the RP group differed from chance level of 8.3% ($p < .001$, Cohen's $d = 0.90$).

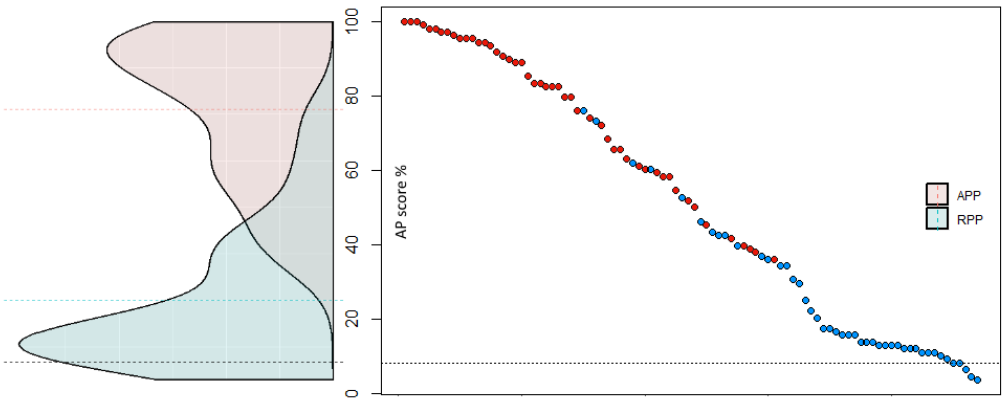

Figure 2.1: Distribution of AP scores (red: AP group, blue: RP group). The black dashed line represents the chance level (8.3%). On the left side, the distributions of APP and RPP are displayed with the corresponding dashed lines representing the means. On the right side, AP scores of all participants are depicted separately.

2.3.3 Task

The EEG experiment consisted of four blocks (Fig. 2) each of them lasting 45 seconds. During two blocks 30 tuned and 30 mistuned tones were randomly presented. Alternating to the tone blocks, two additional blocks with different continuous noise stimuli were presented for 45 seconds. These noise blocks were used to mask the memory trace between the tone blocks. We used two different noise stimuli to avoid boredom and habituation. Furthermore, the four blocks formed a presentation unit that was repeated four times during the experiment. The order of the blocks was randomized but each presentation unit always started with a tone block. Furthermore, after each block an interval of 30 seconds silence was introduced with longer silence (one minute) at the end of each presentation unit in order to provide the participants the possibility to have a rest. During the task, participants were instructed to attentively listen to the auditory stimuli. Thereby, we explicitly refrained from asking the participants to label the tones or to do some other cognitive tasks. The reason was that we wanted to keep tone presentation as natural as possible. In fact, when musicians hear tones they are normally not required to accomplish a given task. In addition, when RP musicians are

instructed to label tones although they are not able to do that, they might use specific cognitive strategies. These strategies might be reflected in idiosyncratic neural activations that might hamper the detection of typical neural activations. During the whole procedure, participants were instructed to look at a fixation cross presented in the middle of the screen. Additionally, three minutes eyes open resting state was recorded.

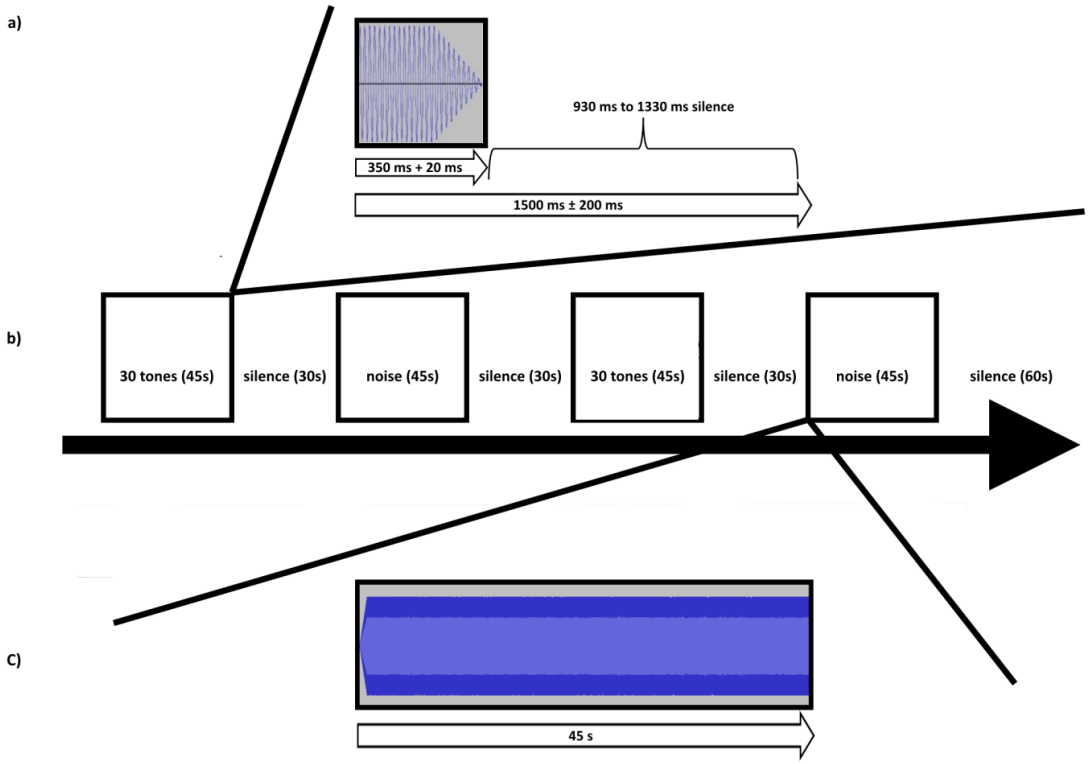

Figure 2.2: Experimental design. a) Representation of a single tone trial. b) Sequence of a whole presentation unit. c) Continuous noise block.

2.3.4 Stimuli

The auditory stimuli consisted of sine tones of different frequencies that were randomly presented, ranging from C4 to B4. Tones were either in tune or slightly mistuned. The mistuned tones were either a quarter halftone flat or sharp respective to the in tune ones. Tone blocks comprised randomly presented tones of different frequencies with a

duration of 370 ms (linear fade-out = 20 ms). The tones were presented with an inter-trial interval of 1,500 ms and a latency jitter of ±200 ms resulting in an inter-stimulus interval in the range of 930-1,130 ms. Additionally, two noise stimuli (Brownian and white) were randomly presented, whereby each kind of noise was presented once in each presentation unit. The noise stimuli were faded in for 20 ms and were continuously presented during 45 seconds. All stimuli were generated and processed with Audacity (Version 1.3 Beta, The Audacity Team, USA).

2.3.5 EEG RECORDING AND PRE-PROCESSING

EEG was acquired using HydroCel Geodesic Sensor Net with 128 Ag/AgCl electrodes, a Net Amp 400 amplifier system, and Net Station acquisition software (Version 5.2.0.2) manufactured by Electrical Geodesic (Electrical Geodesics Inc., Eugene, Oregon, USA). Impedances were kept below 25 kΩ using a potassium-chloride solution. The EEG signal was online-filtered from 0.01 to 100 Hz with an analogue bandpass filter and digitized at a sampling frequency of 1,000 Hz. The vertex electrode Cz served as an online reference. Auditory stimuli were presented binaurally at a sound pressure level of 75 dB using Bose Companion 2 Series III external speakers (Bose Corporation, Framingham, Massachusetts, USA). Participants were asked to relax during the entire procedure to avoid muscle artefacts. Furthermore, prior to the EEG measurement, participants were instructed about the consequences of eye blinks and saccades.

EEG recordings were converted into Brain Vision Analyzer format using EEGLAB (Delorme & Makeig, 2004). Pre-processing was done with Brain Vision Analyzer (Version 2.1, Brainproducts, Germany). As a first step, the outer ring of electrodes was removed due to potential contamination of muscle artefacts (108 electrode sites remained). The data were offline-filtered from 0.1–30 Hz with an infinite impulse response filter. A semi-automatic individual independent component analysis was performed to remove eye movement artefacts (Jung et al., 2000; Li et al., 2006; Vigário, 1997). Afterward, the average reference was calculated and the vertex electrode was re-

used as an active electrode. Additionally, an automatic raw-data inspection was performed to exclude segments with excessive amplitudes. Thresholds were set as follows: 50 µV/ms maximal allowed voltage steps between two sample points, 100 µV maximal allowed absolute difference during a time window of 200 ms, ± 100 µV maximal/minimal allowed amplitude and 0.5 µV lowest allowed activity during an interval length of 100 ms.

2.3.6 ERP ANALYSES

The EEG data related to the tones were segmented corresponding to the different conditions, namely in tune (IT) and mistuned (MT). A baseline correction of -100 ms to stimulus onset was performed. Afterward, data were averaged for each participant and condition, exported, and further analysed with an in-house Matlab script (Version R2015b, Mathworks, USA). This script extracted the signed area (Luck, 2014) over a predefined time window. For the negative components (N100, N200, N400) the signed area underneath the baseline and for the positive component (P200) the signed area above the baseline was exported. One of the main advantages of using the signed area is that fairly large time windows can be used without voltage cancelation and with less bias towards choosing arbitrary time windows. Furthermore, the ERP components are treated as signals that extend over time instead of focusing on the peak voltage or on small time windows. Grand averages were computed separately for both groups (AP, RP) and conditions (IT, MT) (Fig. 3). Furthermore, electrode sites for statistical analyses were selected based on the according to voltage maxima in the grand averages. For the N100 component, signed areas were extracted during a time window of 50–200 ms, whereas for the P200 component area was evaluated in the range of 120–400 ms (Finke et al., 2016). Both components were exported for the Cz electrode site according to the maxima shown in Figure 3. Signed area values corresponding to the N200 and N400 components were extracted in time windows of 270–350 and 350–470 ms at electrode Fz according to the maxima.

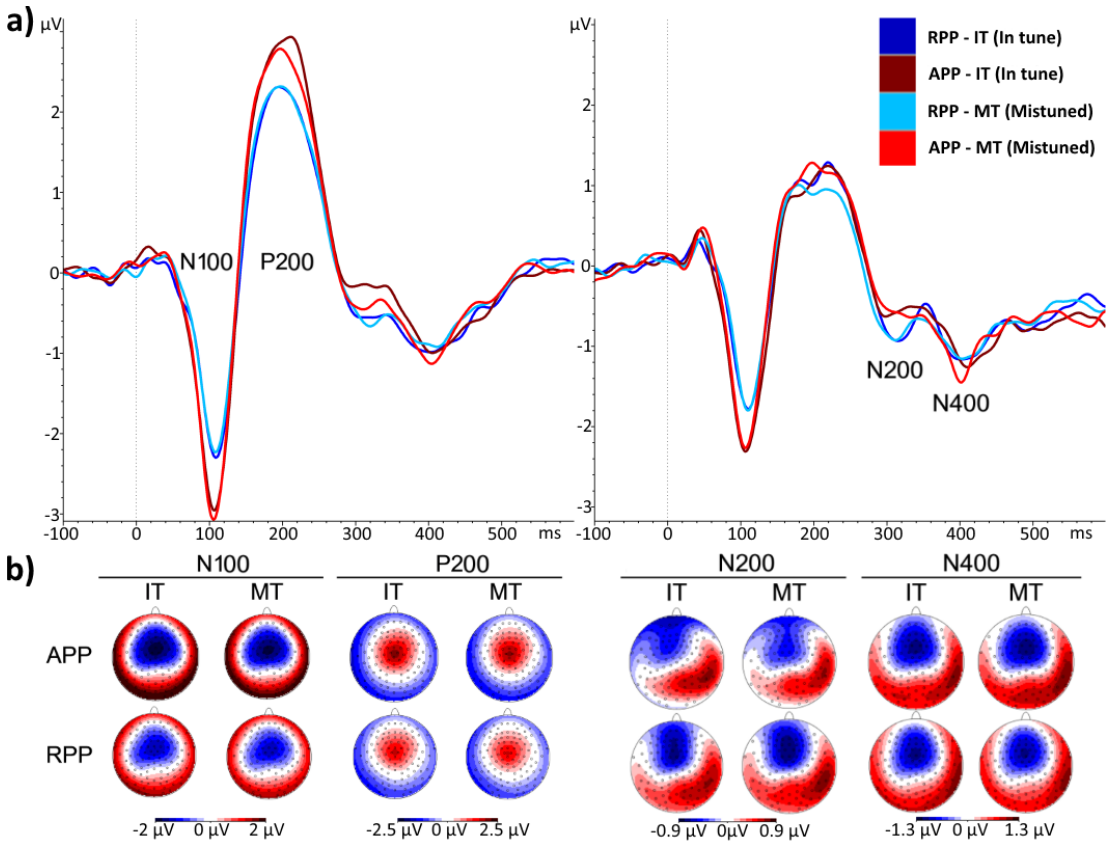

Figure 2.3: ERPs and topographic scalp maps for the different components of interest. ERPs for the APPs are coloured in red and those for the RPPs in blue. Darker colours represent the IT condition and lighter colours the MT condition. a) On the left, ERPs for the N100 and P200 components for electrode site Cz are shown. N200 and N400 at electrode site Fz are depicted on the right. b) Topographic scalp maps for the components of interest and the different conditions (IT, MT) with the corresponding μV scales.

2.3.7 Source reconstruction

For all ERP components that revealed differences in the conventional ERP analyses described above, intracortical sources were computed using the Brainstorm toolbox (Tadel et al., 2011). Based on the fact that we revealed between-group differences that were independent of the experimental conditions (IT, MT), brain responses to all tones were used for source reconstruction. For source reconstruction in each participant, we

used the average files containing all tones. Noise covariance matrices were computed based on individual resting states. Open MEEG software (Gramfort et al., 2010; Kybic et al., 2005) was used to construct head models with 15,002 dipoles (1 x 1 x 1 mm surface grid). This software reconstructs a forward model that best fits brain activity measured at the sensor level. Electrode positions were defined by a standard EGI template provided by Brainstorm. The default anatomy template of the International Consortium for Brain Mapping (ICBM-152) provided by Brainstorm was used. Conductivity values (scalp 0.44, skull 0.018, and brain 0.25) were chosen based on the recommendations from Song et al. (2015). Inverse solutions were computed by using minimum norm imaging to calculate current density maps with constrained source orientations. Based on the voltage measured at the sensor level, inverse solutions estimate the corresponding distributions of active cortical sources.

Before starting with source reconstructions of the single ERP components, the reconstruction of the N100 source (peak) was used as a localizer in order to provide evidence for a meaningful inverse solution. Based on the fact that it has frequently been shown that the main generators of the N100 component are located in the primary and secondary auditory cortex (Jancke et al., 2006; Ott et al., 2011; Pantev et al., 1995; Zaehle et al., 2004), we tested whether this was also true for our data. Figure 4 shows, that this was indeed the case.

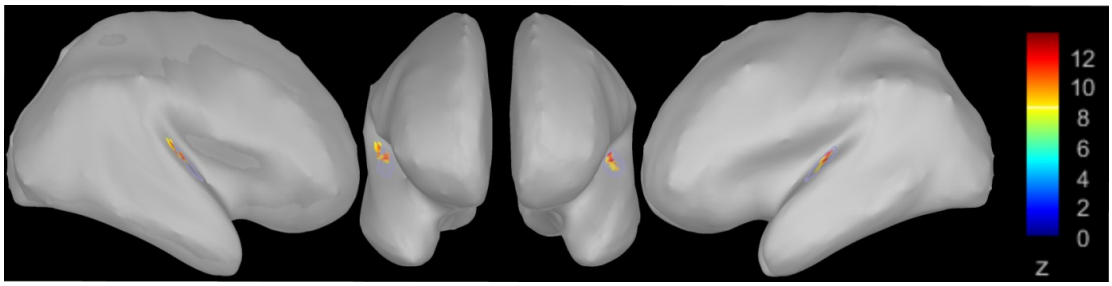

Figure 2.4: Source estimation for the peak amplitudes of the N100 component at 107 ms. Centroid voxels are located in the bilateral Heschl's gyrus (MNI coordinates for the left hemisphere: x = -42, y = -23, z = 13; right: x = 43, y = -24, z = 13). Spatial smoothing was applied for visualization.

Time windows for statistical source analyses of the ERP components were chosen based on the global field power of the grand averages which were computed with Brain Vision Analyzer (Fig. 5). Using the same time windows as for the ERP analyses (i.e., signed area) was not an option because the time windows for the signed area were fairly wide. If the same time windows would have been taken for source reconstruction, the estimation would be contaminated with other neighbour components. To get a good temporal approximation of the single components, global field power was used instead (Skrandies, 1989). Therefore, a grand average based on the segments of all tones was used for the determination of the time windows for source analysis (Fig. 5). Time windows used for source reconstructions were 70–145 ms for the N100 and 145–275 ms for the P200 component.

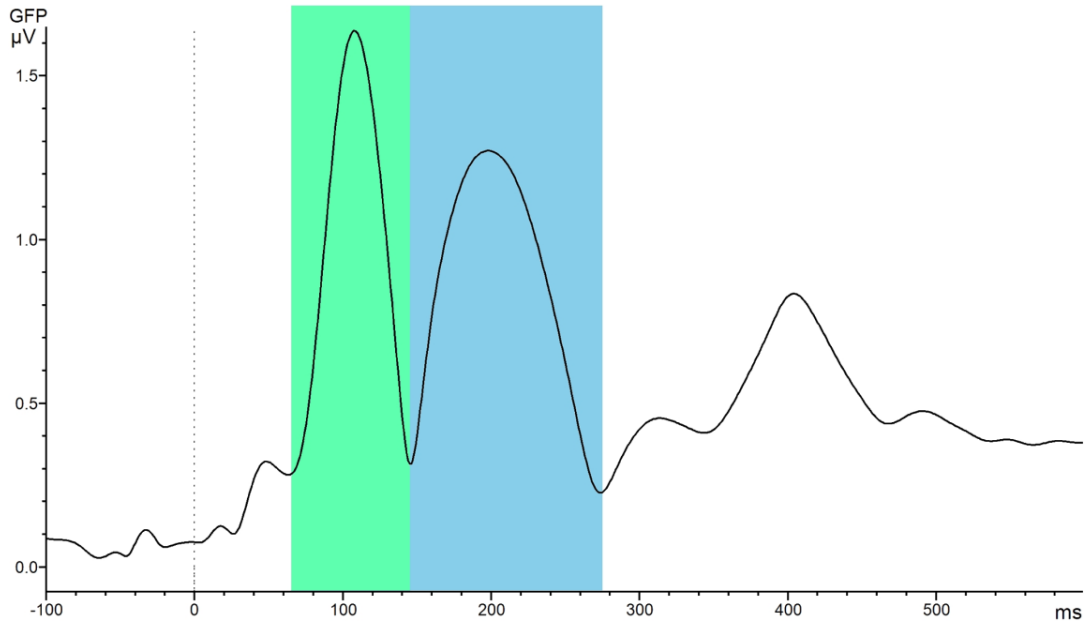

Figure 2.5: Time windows used for source reconstruction based on the global field power for the N100 (70–145 ms, green) and the P200 (145–275 ms, blue) components.

At the single-subject level, z-score normalizations were performed relative to baseline (-100 ms to stimulus onset) to make amplitudes comparable between subjects. Furthermore, a priori defined ROIs containing temporal, frontal, and parietal areas were chosen based on previous literature on AP (Bermudez et al., 2009; Elmer et al., 2013,

2015; Hirata et al., 1999; Jäncke et al., 2012; Ohnishi et al., 2001; Schneider et al., 2005; Schulze et al., 2011; Schulze et al., 2009; Wengenroth et al., 2014; Zatorre et al., 1998; Zatorre, 2003). Figure 6 shows the 18 ROIs, which were selected based on the Desikan-Killany atlas (Desikan et al., 2006). There was one source signal for each vertex (e.g., voxel) contained in a specific ROI because constrained dipole orientations were used. Furthermore, grand averages of the source reconstructions for each group were calculated, as well as differences between the groups (AP-RP) to visualize which group had stronger/weaker activations.

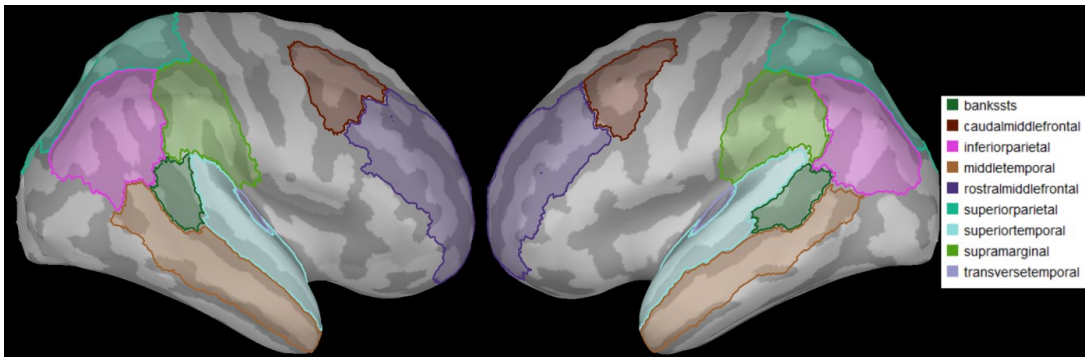

Figure 2.6: Nine ROIs were selected in each hemisphere. The names of the ROIs are depicted on the right with the corresponding colours.

2.3.8 Statistical analysis

ERP

Four two-way (2x2) ANOVAs with repeated measurements on one factor (IT vs. MT) and one grouping factor (AP vs. RP) were computed with area values separately for each component of interest (N100, P200, N200, N400). Before computation, we tested whether the data were normally distributed. Statistical analyses were executed with SPSS (Version 22, IBM, USA). As dependent variables, we used signed area values measured at electrode Cz for the N100 and P200 components and at electrode Fz for the N200 and N400 waveforms.

Post-hoc bivariate Pearson correlations (one-tailed) were calculated between AP scores and signed area values for each ERP component in the whole sample of participants as well as separately for the two groups. Since the correlations served for descriptive purposes, no correction for multiple comparisons was applied.

Source reconstructions

FieldTrip source statistics (Oostenveld et al., 2011) were performed calculating independent t-tests (two-tailed) for the intracortical sources in the predefined ROIs for each ERP component that revealed a significant result in the previous analyses. In particular, we performed non-parametric cluster-based permutation statistics with an alpha set at $p = 0.05$. Here, samples exceeding a t-value associated with an alpha of 0.05 were clustered according to their spatial adjacency. Each cluster's sum of the t-values was the basis for its cluster-level statistic, and group differences were tested with the maximum of those statistics. By setting the threshold for the cluster alpha to $p = 0.05$, a t-value thresholding at the 2.5th and the 97.5th quantiles was attained. Significance probability was calculated with the Monte Carlo method using 50'000 randomly selected permutations. A cluster-wise correction for multiple comparisons was applied.

Furthermore, bivariate Pearson correlations (one-sided) were computed for the whole sample and also separately for the two groups between AP scores and the current density values of the ROIs that revealed a significant group difference in the cluster-based permutation test. For each of these ROIs, activity was averaged over all vertices included in the corresponding ROI. Furthermore, an average over time was calculated for the duration of the whole component. This resulted in one value per ROI and participant.

2.4.1 RESULTS WITH RESPECT TO THE AMPLITUDES OF THE N100, P200, N200 AND N400 COMPONENTS

As depicted in Figure 7, the evaluation of the N100 component revealed a main effect of group ($F(1,92) = 4.801$, $p = .031$, $d = 0.45$, $CI_{95\%} = 0.043–0.863$). The AP group showed larger area values compared to the RP group, irrespective of the condition (IT, MT). No other main effect or interaction reached statistical significance. For the P200 component there was a tendency towards a significant group difference ($F(1,92) = 3.574$, $p = .062$, $d = 0.391$, $CI_{95\%} = 0.018–0.799$) with APPs showing slightly larger signed areas in comparison to RPPs. For the N200, and N400 components we did not reveal significant results.

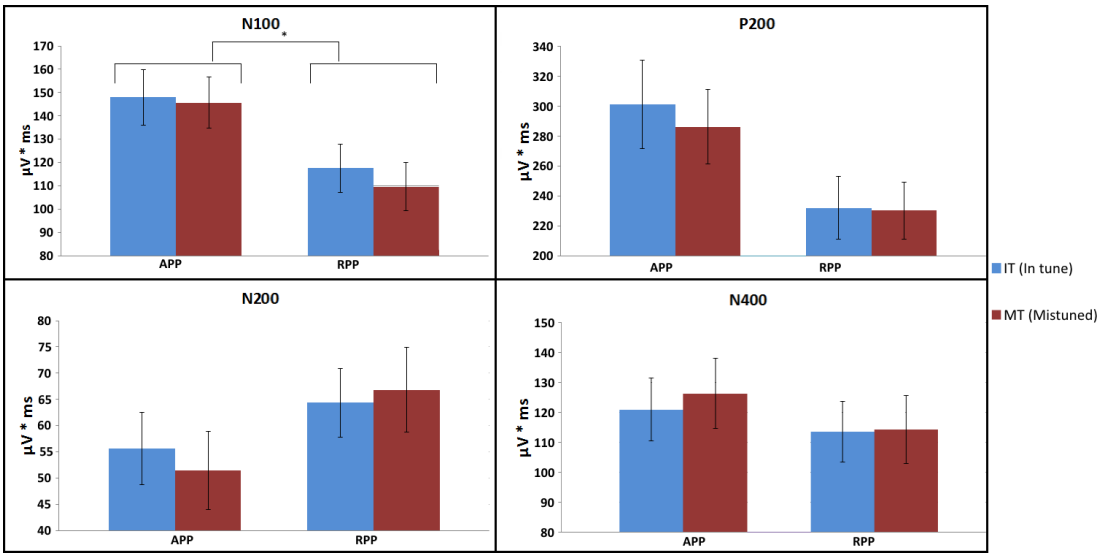

Figure 2.7: Signed area values for the components of interest (N100, P200, N200, and N400). Means and standard errors of means are displayed. Values for the AP group (APP) are shown on the left while those for the RP group (RPP) are displayed on the right. Furthermore, amplitudes for the in tune (IT) condition are blue coloured while those for the mistuned (MT) condition are red. The unit for the signed area is µV*ms.

Weak but significant correlations (uncorrected) were only found for the P200 component. In particular, within the AP group we revealed a positive relationship between AP score and the MT-related area values (r = .24, p = .047). Otherwise, within the RP group, we found a negative correlation between AP score and P200 area (r = -.28, p = .032; Table 3).

Table 2.3: Correlations between the AP scores and ERP area depicted for all components and conditions (IT: in tune, MT: mistuned). The correlation coefficient r and the p-value (uncorrected) are displayed for all participants (All), and also for the two groups (AP and RP group) separately.

		IT N100 **MT**		**IT** P200 **MT**		**IT** N200 **MT**		**IT** N400 **MT**	
All									
	r	.19	.24	.18	.18	.00	-.08	.11	.15
	p	.032*	.009**	.045*	.042*	.486	.210	.150	.079
AP group									
	r	.12	.19	.22	.24	.22	.21	.05	.14
	p	.207	.097	.063	.047*	.063	.079	.358	.165
RP group									
	r	-.00	-.05	-.28	-.23	.02	-.12	.19	.15
	p	.495	.365	.032*	.062	.444	.217	.110	.169

2.4.2 Results with respect to the intracortical sources of the N100 and P200 components

Based on the fact that the ERP analyses revealed a main effect of group for the N100 component and a tendency towards significance for the P200 component, source reconstructions were only calculated for these two components. For the N100 component (70–145 ms) significant clusters between AP and RP groups were found in

the right hemisphere (Fig. 8). Between-group differences were most pronounced in three ROIs, namely STS (cluster size = 68, p = .002), middle temporal gyrus (MTG) (cluster size = 58, p = .027), and the transverse temporal cortex, which corresponds to Heschl's gyrus (HG) (cluster size = 72, p = .004). Since permutation statistics do not enable to determine the direction of the effects, we additionally computed between-group differences of the source activations (AP-RP) (Fig. 8). No significant result was revealed for the P200 component.

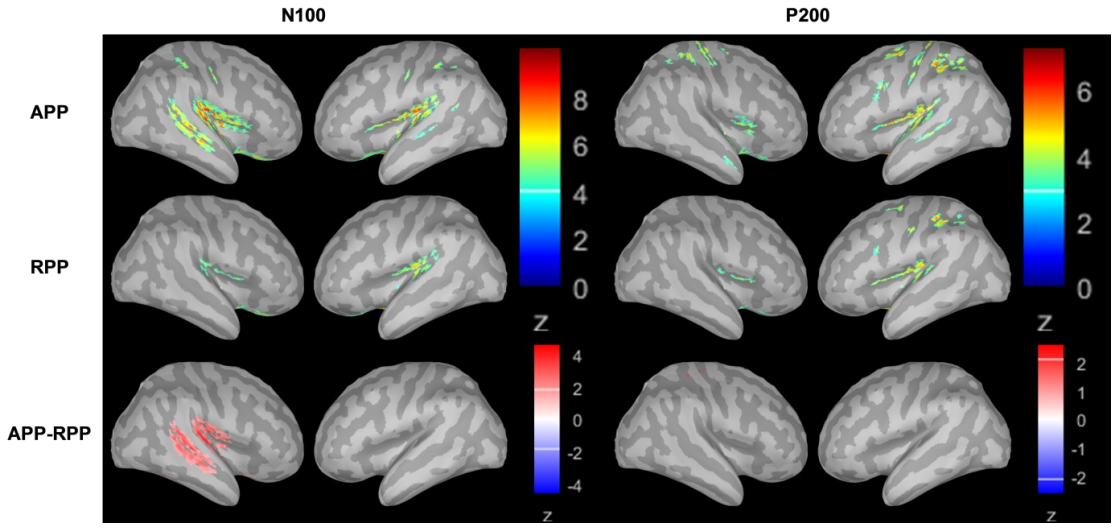

Figure 2.8: Source reconstructions in the AP and RP groups for the N100 and P200 components (z-transformed relative to the baseline). Source activations for the N100 are depicted on the left side, those for the P200 on the right side. The upper row shows the source activation for the AP group and the middle row for the RP group. In the bottom row, the differences between the source activations of the AP group and the RP group are displayed (AP group-RP group). Z-scores in the bottom row were thresholded at z = 2.1, which corresponds to a p-value of .01. Activations in the STS, MTG, and HG in the right hemisphere showed statistical differences between the two groups. Spatial smoothing was applied for visualization purposes.

The results of the correlations between the AP scores and the ROIs showing group differences are displayed in Table 4. These additional analyses revealed a significant correlation for the right MTG in the AP group (r = .33, p = .010) (Fig. 9).

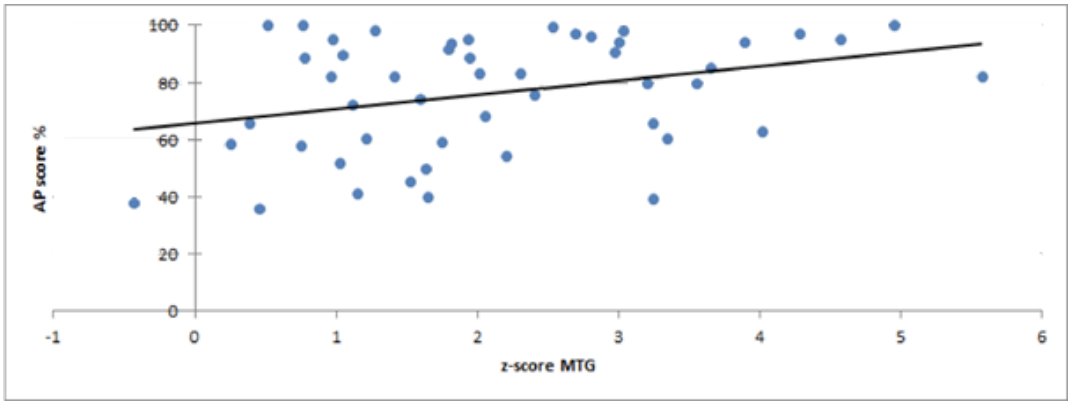

Figure 2.9: Correlation between the AP score and the z transformed activation in the middle temporal gyrus (MTG) for the AP group.

2.5 Discussion

2.5.1 General discussion

In the present study, we examined the neural underpinnings of AP processing during attentive tone listening. We explicitly used attentive listening without an additional cognitive task (e.g., tone labelling or tone discrimination, etc.) in order to keep the listening condition as ecologically valid as possible. In normal music and auditory listening situations, musicians are not explicitly asked to solve a task related to the listening condition. They simply listen and apply the analysis strategies, which they have largely implicitly access to. In the case that AP musicians use specific auditory object analysis strategies automatically, we aim to identify the neural underpinnings of these strategies with this study. To achieve this, we used two different approaches: First, we analysed the N100, P200, N200, and N400 components to conform to earlier ERP studies on AP musicians. Our study was based on the so-called two-component model of AP processing (Levitin, 1994) as well as on studies assuming an early perceptual categorization (or perceptual memory) and a subsequent cognitive labelling process (Elmer et al., 2013, 2015). Differences in perceptual categorization between both groups should thus be reflected in differences in early ERP manifestations, namely the N100 and P200 components and their associated intracortical sources. In the case of

significant differences for the subsequent labelling process, we anticipated significant differences for the late ERP components and their associated intracortical sources, namely the N200 and N400 components. Significant group differences were only identified for the N100 (and partially for the P200), with APPs exhibiting larger area values for both components. Second, we analysed the intracortical sources of the N100 component and revealed stronger current densities in APPs in the right STS, MTG, and HG. In addition, we observed a weak but significant correlation between activation in the MTG and AP performance for the AP group.

In our work, we also hypothesized that only AP musicians would recognize the slightly mistuned tones and therefore exhibit different neuronal responses compared to RPPs. We expected this difference in any component of interest (N100, P200, N200, and N400). This should be reflected by an interaction effect of the two-way repeated measurement ANOVAs performed with the area values. However, no such interaction effect was found. After the EEG experiment, participants completed a short questionnaire to evaluate whether the mistuned tones were identified as such. The question was whether the participants consciously noticed that some tones were mistuned (answer alternatives: yes/no). APPs (45 of 49), as well as RPPs (34 of 45), mentioned that they detected the mistuned tones. Furthermore, RPPs and APPs did not significantly differ in terms of musical experience. Therefore, it is possible that this could account for the missing interaction effects. In the following section, we will discuss how and whether our findings fit the existing literature.

2.5.2 Discussion with respect to the ERP components

Studies examining differences between AP and RP musicians with respect to the N100 amplitude during tone listening paradigms are relatively rare (Behroozmand et al., 2014; Elmer et al., 2013; Hirata et al., 1999; Hirose et al., 2003, 2005; Hirose et al., 2004; Itoh et al., 2005; Pantev et al., 1998; Wu et al., 2008). Such studies have applied different paradigms (passive listening, attentive listening, tone labelling, tone discrimination, and

vocal feedback discrimination), and examined different samples (children and adults) with mostly small sample sizes (7-22 AP musicians). Since these studies differed in many aspects, the reported results with respect to the N100 component are undoubtedly mixed. The majority of these studies have reported little or no differences between AP and RP musicians. One of the first studies of this type did not observe any differences in the magnetic counterpart of the N100 (N100m) component between APPs and RPPs (Pantev et al., 1998). In this study, participants had to listen to piano and pure tones while watching a cartoon. The authors, however, identified enlarged dipole moments in the auditory cortex in response to the piano tones in comparison to sine tones. This enlargement was generally present in musicians, with no differences between APPs and RPPs. In a follow-up study of the same research group, different activations for the N100m component between AP and non-musicians were reported (Hirata et al., 1999). The participants of this study had to listen to different stimuli while counting some of them to ensure attentional processing of the tones. Hirose et al. (2003, 2005; 2004) used tone identification and passive listening tasks and reported stronger N100m responses to tones among AP musicians. These stronger N100m amplitudes were also associated with generally stronger dipoles bilaterally in the auditory cortex. Further evidence for differences during early pitch processing stages comes from Itoh et al. (2005), who identified a negativity over left posterior-temporal electrode positions, an ERP component taking place around 150 ms after stimulus onset and which was only present in AP musicians performing excellently in an AP task (High-AP). This component (AP negativity) was present in both listening and pitch-labelling tasks. Furthermore, Wu et al. (2008) recorded conventional ERPs to auditory stimuli during three experimental conditions (inattentive listening to tones, tone labelling, and relative pitch labelling), and identified no between-group differences for all conditions. However, analysis of the intracortical sources of ERPs uncovered stronger and more widespread activations bilaterally in the auditory cortex as well as in the occipital and parietal areas, particularly during the tone labelling task. Another study from Elmer and colleagues (2013) using a passive listening paradigm reported no difference between AP and RP musicians with respect to the N100 amplitude. The most recent study (Behroozmand et

al., 2014) using a vocal pitch error detection task uncovered a massively stronger N100 response to pitch-shifted voice feedback over right-sided electrode clusters for both AP and RP musicians compared to the non-musician group. The P200 component obtained over a left-sided electrode cluster was increased in AP musicians compared to RP musicians.

Taken together, the aforementioned studies have reported mixed results with respect to N100 or N100m amplitudes. Previous studies reported enlarged N100 amplitudes to musical stimuli during tone identification (Behroozmand et al., 2014; Hirata et al., 1999; Hirose et al., 2003, 2005; Hirose et al., 2004). However, it is difficult to compare these studies due to their differences with respect to the paradigms used, samples, and analysis techniques. However, it cannot be ruled out that attention or other cognitive processes might have influenced the results since the paradigms require different levels of attention and stimulate different cognitive processes. This is particularly problematic for RP musicians, as some of the required tasks (e.g., labelling) are difficult or even impossible for them. However, when AP musicians automatically use a particular processing strategy, they should have used it particularly during attentive listening. In fact, we observed enlarged N100 components in AP musicians, a result which supports the notion that early auditory processing steps are different in AP musicians. Thus, AP musicians either allocate more attention to the tones or automatically categorize the tones at early processing steps. Nevertheless, we have to be careful not to over-interpret this finding since the effect size for this between-group differences was small to medium (Cohen's d = 0.453) and most importantly much smaller than the effect size observed for the performance in the AP test (Cohen's d = 2.66). Thus, the N100 amplitude did not reflect all neurophysiological processes involved in AP processing. However, the exact psychological functions responsible for this are difficult to determine. This problem will be explained in greater detail in the following section, which discusses the intracortical sources of the N100 component, as these results help us understand the neural underpinnings of AP processing.

2.5.3 Discussion on the intracortical sources of the N100 component

The intracortical sources of the N100 component revealed substantial between-group differences with respect to current densities in the right-sided STS, MTG, and HG. These brain areas are part of the perisylvian brain, which is distinctly organized anatomically and functionally in AP musicians compared to RP musicians and non-musicians (Hänggi et al., 2008; Jäncke et al., 2012; Keenan et al., 2001; Kim & Knösche, 2016; Loui et al., 2012; Loui et al., 2011, 2012; Oechslin et al., 2010; Ohnishi et al., 2001; Schlaug et al., 1995; Schulze et al., 2013; Schulze et al., 2009; Wengenroth et al., 2014). Based on previous studies, it is not entirely clear whether the right, left, or both perisylvian areas are differently involved in AP processing. Some anatomical studies reported specific anatomical features on the right side (Wengenroth et al., 2014), while others have identified a stronger relative leftward asymmetry (Schlaug et al., 1995). The same pertains to neurophysiological activations obtained from fMRI, EEG, and MEG studies (Gaab et al., 2006; Itoh et al., 2005; Ohnishi et al., 2001; Schulze et al., 2013; Schulze et al., 2009). Based on the current status of knowledge, it is not easy to reconcile these different findings, since most of the studies (especially the functional studies) differ in terms of the applied task, statistical analyses, and sample size. Therefore, it remains challenging to use the available literature to reconcile the right-sided activation preponderance in the STS, MTG, and HG that occurs during the first 100 ms of auditory processing. Otherwise, the increased brain activity we revealed in APPs in the right HG, STS, and MTG might be reconcilable with previous models of auditory processing suggesting an advantage of right-sided perisylvian areas for spectral processing (Poeppel, 2003; Zatorre et al., 2002). Furthermore, it is noteworthy to mention that the tones we used in our experiment had a duration of 370 ms, which roughly corresponds to a frequency of 3 Hz. Low-frequency oscillations have previously been shown to be more pronounced in the right compared to the left auditory cortex (Giraud et al., 2007), and to play an important role in packing the multi-time speech signal into units of the appropriate temporal granularity (Giraud & Poeppel, 2012).

Previous fMRI and PET studies on AP musicians have used pitch memory tasks (Gaab et al., 2006; Schulze et al., 2009), auditory Stroop tasks (Schulze et al., 2013; Zatorre et al., 1998), pitch detection tasks (Schulze et al., 2013; Zatorre et al., 1998), or simply passive listening to a musical piece as stimulation (Ohnishi et al., 2001). In addition, prior studies have largely worked with small samples only including a few AP musicians (10-18). Thus, these studies can hardly be compared to our work, which explicitly used an attentive listening paradigm to keep the listening condition as ecologically valid as possible. This is important in that musicians are normally not asked to label, discriminate, or rehearse single tones in natural music listening conditions but simply play and listen to music. In addition, we included a large sample of 49 AP musicians in order to avoid false positive results. Nevertheless, besides the discrepancies observed in the results and paradigms used, some similarities exist between our findings and the results of the aforementioned studies.

In line with several previous studies, we assume that AP musicians (in contrast to RP musicians) automatically categorize tones during very early processing stages, even when they are not required to do so. In these early processing stages (approximately 100 ms after stimulus presentation), many bilateral perisylvian brain areas—particularly the HG, STG, MTG, and STS—are involved in a slight preponderance of the right-sided areas (see Fig. 8). Several studies have shown that the STS, MTG, and STG are brain areas that are involved in categorization and multimodal integration processes. For example, Klein and Zatorre (2011) showed larger right- and left-sided STS activity during categorical tone processing in non-AP subjects. Moreover, left-sided involvement of the STS has been reported by Schulze et al. (Schulze et al., 2013; Schulze et al., 2009) in pitch memory and auditory Stroop tasks for AP musicians compared to RP musicians. They interpreted these findings as evidence for an early encoding process which is involved in the categorization of tonal information into pitch chroma classes and is most likely controlled by the STS. Activations in the MTG have previously been associated with multimodal processing (Zatorre et al., 1998), lexical processing (Hickok & Poeppel, 2007; Oechslin et al., 2010), and access to stored pitch

categories (Loui et al., 2011, 2012). Therefore, it might be that the MTG is functionally stronger associated with tone categorization in AP musicians, while the right HG might generally play an important role in AP, as highlighted by Wengenroth et al. (2014). Additionally, Wengenroth and colleagues observed increased volumes of the HG in APPs, which was highly correlated with AP performance. Accordingly, the authors suggested that the right HG is a pivotal structure for AP perception, and stated that the left hemisphere is important for the labelling process.

In the present study, the preponderance of right-sided activation during the very early auditory processing stages could indicate a process specific to AP musicians. We assume that this specific process reflects automatic tone categorization that is supported by a network comprising several perisylvian brain areas. This network, which is lateralized to the right hemisphere, is different from what is known from studies examining neural activations in these brain areas during different attention conditions (Jancke et al., 2001; Jäncke et al., 1999). When attention is directed to both ears, hemodynamic responses increase bilaterally in the perisylvian brain, and not only in one hemisphere. Thus, we assume that the right-sided activation increase identified in the present study indicates an AP-specific neural activation, which is most likely related to automatic tone categorization.

In this context, we would like to emphasize that several EEG experiments have shown that brain activation can quickly change between the right and the left perisylvian brain, even during the first 1,000 ms of auditory information processing (Baumann et al., 2008; Sinai & Pratt, 2003). Thus, lateralized activations obtained from fMRI and EEG/MEG experiments are difficult or even impossible to compare since both techniques measure neurophysiological processes at different time scales. Finally, it is important to mention that according to our results we did not reveal functional evidence reflecting pitch labelling mechanisms. In fact, pitch labelling has more likely been associated with late EEG manifestations (Elmer et al., 2013) as well as with activation patterns in the language-dominant left hemisphere (Deutsch et al., 2004; Wu et al.,

2008; Zatorre et al., 1998). Accordingly, we propose that the increased current density values we revealed in the right hemisphere in AP compared to non-AP musicians reflected an optimization of auditory objects recognition or categorization processes rather than labelling per se (Kim & Knösche, 2016).

2.5.4 Possible implications for AP perception and labelling

Our study indicates that AP musicians automatically activate a right-sided perisylvian brain network during attentive tone listening. This additional activation might be the neural underpinning of tone classification. However, it is also conceivable that AP musicians allocate more attention to the incoming tone stimuli, thereby resulting in a more widespread activation in the perisylvian brain. However, even though one may assume that attentional enhancement and tone categorization interact when AP musicians listen attentively to tones, attentional enhancement effects could not entirely explain this additional right-sided activation since attentional enhancement effects are mostly bilateral (Jäncke et al., 1999, 2001; Noesselt et al., 2002). In addition, it is noteworthy to mention that we did not reveal between-group differences in the MT condition, leading to suggest that attention functions cannot substantially explain the increased right-sided activity we revealed during early stages of auditory processing in AP compared to non-AP participants.

Interestingly, we did not find between-group differences for later processing stages, which are thought to control tone labelling. This result is unexpected since tone labelling has been proposed to be the idiosyncratic trait of AP (Levitin & Rogers, 2005). A possible explanation could be that the labelling process occurred very early in time and involved perisylvian brain areas. Otherwise, it is also possible that the labelling process itself is not fully automatic as hitherto assumed but only activated when requested. Finally, it is important to remark that previous authors suggested a contribution of the prefrontal cortex to labelling processes (Bermudez & Zatorre, 2005; Ohnishi et al., 2001; Zatorre et al., 1998). However, signal transmission from the

auditory cortex to anterior brain sites needs time and is therefore expected to be reflected by other ERP components than the N100 response. This perspective is also compatible with the results of the source estimation we used showing that the N100 was associated with current densities in auditory-related brain areas, which have previously not explicitly been associated with labelling processes.

2.6 Limitations and Outlook

A limitation which concerns AP studies, in general, is how to measure and objectify this ability. Many studies (as we did) rely on self-report for classifying musicians into AP and RP musicians. The reason for this procedure is that it is not entirely clear which performance cut-off should be used to decide whether the musicians are AP or RP musicians. A more objective and widely accepted measurement is urgently needed for this field (e.g. mean absolute deviation as proposed by Bermudez and Zatorre (2009)) because the answers of the participants may also rely on self-esteem, social desirability, or other motivational reasons. Some people possess AP for a single tone and when asked to identify a given tone they can do it by calculating the interval between those two tones. This kind of quasi-AP (Takeuchi & Hulse, 1993) is a strategy which needs more time than simply identifying a tone without a reference tone. Therefore, exact reaction time measurements could be useful for evaluating AP abilities (Bermudez & Zatorre, 2009). In addition, future studies should try to develop more sensitive tests that enable a meticulous estimation of both AP and RP abilities. This implies that instead of merely improving the sensitivity of AP tests it would be advantageous to additionally implement a test that screens RP abilities. Furthermore, to investigate whether early and/or later processing stages are involved in determining AP ability, more studies that investigate basic auditory processing are needed. In this context, it is very important to also publish those studies with null effects. Although we explicitly conducted this study using attentive listening and not labelling, it will be interesting to more carefully examine the neurophysiological underpinnings of tone labelling, tone discrimination, and pitch memory in AP and RP musicians. Finally, one might argue that we should

have used subgroups of AP musicians in order to examine whether different performance levels in the AP test might be related to the measured neurophysiological measurements. We refrained from subdividing the AP group into several subgroups in order to keep the AP group large enough for sustained enough statistical power. However, we performed several correlation analyses to examine whether the AP performance correlated with the amplitudes of the ERP components and the current densities values of perisylvian brain areas. These correlations were mostly small, indicating that the neurophysiological responses to attentive tone listening do not entirely depend on the performance in the standardized AP test.

2.7 Conclusion

The present work included the largest number of AP and RP musicians used in any neuroscientific study on this topic to date and examined whether attentive tone listening evoked different neural responses in AP and RP musicians. We identified differences at the early stages (N100) of auditory processing between AP and RP musicians, which were accompanied by stronger activations in right-sided perisylvian brain areas comprising the STS, MTG, and the HG in APPs. In contrast, no differences were observed at the later stages of auditory processing. The data presented here suggest that differences between AP and RP musicians exist at early stages of auditory processing reflecting categorization, which is pivotal for AP musicians. This early auditory processing is controlled by bilateral perisylvian brain areas with a preponderance of the aforementioned right-sided perisylvian network.

The Importance of the Fibre Tracts Connecting the Planum Temporale in Absolute Pitch Possessors

Anja Burkhard[§*1], Jürgen Hänggi[§1], Stefan Elmer†[1], Lutz Jäncke†[*1,2]

[1] Division Neuropsychology, Institute of Psychology, University of Zurich, 8051 Zurich, Switzerland
[2] University Research Priority Program (URPP), Dynamics of Healthy Aging, University of Zurich, 8051 Zurich, Switzerland

[§]Shared first authorship
†Shared last authorship
[*]Corresponding authors

This article was originally published in:

Burkhard, A., Hänggi, J., Elmer, S., & Jäncke, J. (2020). The Importance of the Fibre Tracts Connecting the Planum Temporale in Absolute Pitch Possessors.
Neuroimage, 211, 116590.
doi: https://doi.org/10.1016/j.neuroimage.2020.116590

3.1 Abstract

In the present study we investigated 90 musicians of both sexes who possess different degrees of absolute pitch (AP) using diffusion tensor imaging in association with a correlational approach and evaluated whether there are differences in perisylvian connections depending on the proficiency level of AP. We expected higher fractional anisotropy with increasing AP ability in the white matter underlying perisylvian brain areas. Results revealed a significant positive correlation between the proficiency level of AP and fractional anisotropy values in the left-sided white matter underlying the planum temporale. We interpret this result as an indicator of auditory processing differences between musicians possessing different degrees of AP, reflecting early auditory encoding and categorization processes. The present study provides further evidence for the substantial importance of the left-sided planum temporale for the ability of AP.

3.2 Keywords

Absolute pitch, DTI, musicians, fractional anisotropy, planum temporale

3.3 Introduction

People with normal hearing abilities are usually able to recognize whether a perceived tone is higher, lower, quieter, or louder compared to another tone that occurs in sufficient temporal nearness. However, some people are able to classify pitches absolutely into tone categories without the use of a reference tone (Bachem, 1937). This relatively rare ability is called absolute pitch (AP). So far, a vast range of AP abilities has been described in previous studies ranging from excellent to random performance with many intermediate levels (Bachem, 1937; Ward et al., 1982). Since a partial AP ability occurs more frequently than a perfect or no AP ability (Wengenroth et al., 2014),

many authors have suggested that the ability of AP is gradually distributed rather than being an all-or-nothing phenomenon (Bermudez & Zatorre, 2009; Itoh et al., 2005; Wilson et al., 2009). However, the neurophysiological and cognitive underpinnings of AP are not entirely understood. In fact, it is still a matter of debate whether early perceptual processes (Burkhard et al., 2019; Hirata et al., 1999; Schulze et al., 2009; Wengenroth et al., 2014) or later cognitive processing stages determine AP (Elmer et al., 2013; Levitin & Rogers, 2005; Zatorre, 2003). The latter are thought to be mostly modulated by frontal brain regions, whereas the former are dependent upon perisylvian areas. In this study, we focused on white matter (WM) underlying perisylvian brain areas, which are thought to be important for perception, encoding, and categorization of tonal information at the initial stage of auditory processing (Schulze et al., 2009; Wengenroth et al., 2014).

One of the most discussed perisylvian brain regions subserving AP is the planum temporale (PT), which is involved in auditory processing and supports categorical perception (Griffiths & Warren, 2002; Keenan et al., 2001; Schlaug et al., 1995). Several studies reported an atypically strong left-sided PT anatomical asymmetry in AP musicians, its cause is discussed controversially in the existing literature (Griffiths & Warren, 2002; Schlaug et al., 1995). Whether this well-documented strong left-sided PT asymmetry might be related to the also reported strong right-sided anatomical asymmetry of Heschl's gyrus (HG) is currently not clear. Nevertheless, these studies document an atypical anatomical of the perisylvian brain. In addition, several studies reported positive correlations between pitch labeling abilities and left-sided PT activation (Wilson et al., 2009; Zatorre et al., 1998). Connectivity studies also showed atypical connectivity patterns in AP compared to relative pitch (RP) musicians. Studies applying diffusion tensor imaging (DTI) protocols found increased fractional anisotropy (FA) values in absolute pitch possessors (APPs) (Dohn et al., 2015; Loui et al., 2011; Oechslin et al., 2009), which are often interpreted as indicators of increased anatomical connectivities. Loui et al. (Loui et al., 2011) revealed increased volumes of WM tracts connecting the posterior superior temporal gyrus (pSTG) with the ipsilateral posterior

middle temporal gyrus (pMTG) for APPs compared to non-APPs, whereby AP performance was predicted by the left-sided tract volume. Moreover, the pSTG as well as the pMTG, are connected to the inferior frontal gyrus through the arcuate fasciculus (Catani et al., 2005), which is part of the superior longitudinal fasciculus (SLF). A previous study showed increased connectivity in the left SLF with higher FA for APPs (Oechslin et al., 2009). Furthermore, (Kim & Knösche, 2016) revealed enhanced intracortical myelination in the right planum polare (PP) for APPs compared to non-APPs.

To emphasize the gradual nature of the AP ability, 90 professional musicians were investigated using a correlational approach. Relationships between the proficiency level of AP and perisylvian WM were assessed using a pitch-labeling score and FA as a measure of WM integrity. Based on the existing literature (Dohn et al., 2015; Loui et al., 2011; Oechslin et al., 2009), we expected that musicians with higher AP proficiency will show increased FA compared to musicians with lower AP proficiency. To our knowledge, this is the first study that investigates FA of APPs by taking into account AP ability as a continuous measure instead of using arbitrary group classifications.

3.4 Materials and Methods

3.4.1 participants

The participants of this study already participated in a recently published electroencephalography (EEG) study of our group (Burkhard et al., 2019). Based on the fact that AP faculty is considered as a continuous rather than a dichotomous ability (Bermudez et al., 2009; Itoh et al., 2005; Levitin & Rogers, 2005; Wilson et al., 2009), DTI data of 103 highly trained musicians with different levels of AP performance were acquired (Table 2). Participants were recruited from music conservatories and orchestras and all were either professional musicians, music students, or highly-trained amateur musicians. Since in the former EEG study nine subjects had to be excluded due

to drug abuse and/or psychiatric disorders (depression and anxiety), these participants were also excluded from DTI analyses. Four additional participants had to be excluded because no DTI data were available, resulting in a total sample size of 90 subjects. All participants gave written informed consent, were paid for their participation and the study procedures were in accordance with the Declaration of Helsinki and the Ethics Commission of the University of Zurich approved the study.

AP ability was tested with a modified version of an in-house AP test (Oechslin et al., 2010) that was administered at home by the participants. During this test, the participants had to label 108 sine tones ranging from C3 to B5 (A4 = 440 Hz) that were presented in a pseudo-randomized order. The pitch-labeling score was defined as the percentage of correct responses. A trial was counted as correct if the exact pitch chroma was identified. Errors in pitch height were neglected because previous studies showed that AP musicians make octave judgment errors occasionally (Miyazaki, 1989). As depicted in Figure 1, AP test performance of the subjects varied from random to perfect.

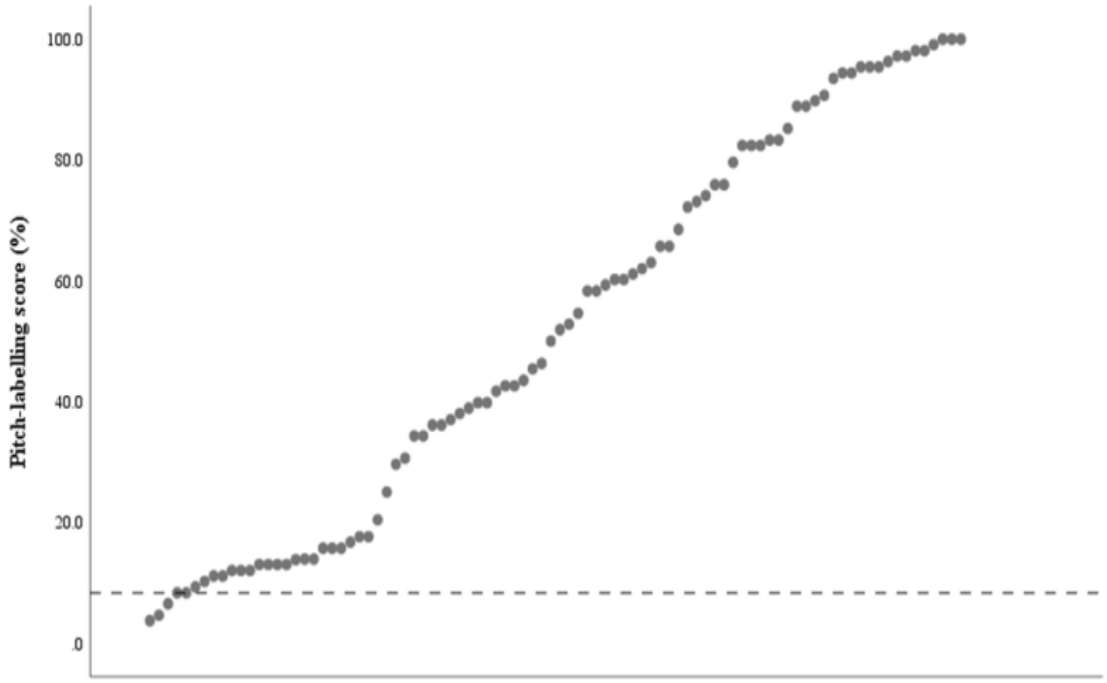

Figure 3.1: Pitch-labeling scores (% correct answers). The chance level is indicated by the dotted line (8.3%).

Musical expertise was assessed using in-house questionnaires, whereas handedness was measured using the Annett questionnaire (Annett, 1970). General cognitive ability was surveyed using a German intelligence test (KAI: Kurztest für allgemeine Intelligenz (Lehrl et al., 1992). Furthermore, the stabilized musical aptitude of the participants was evaluated using the Advanced Measures of Music Audiation (AMMA) test (Gordon, 1989). This test enables to evaluate the ability to discriminate two given melodies by their rhythmic or tonal properties. The test was designed to evaluate musical aptitude regardless of musical achievement. According to pure-tone audiometry, all participants demonstrated a normal audiological status (MAICO Diagnostic GmbH, Berlin, Germany).

3.4.2 DTI DATA ACQUISITION AND PREPROCESSING

MRI data were acquired on a 3T Philips Ingenia scanner (Philips Medical Systems, Best, The Netherlands) using a 15-channel head coil. Single-shot spin-echo echo-planar imaging along 64 different geometrical directions was applied to obtain diffusion-weighted images with a b-value of 1000 s/mm^2 and one reference volume (non-diffusion-weighted). Further scan parameters were echo time = 89 ms, repetition time = 10,022 ms, flip-angle $\alpha = 90°$, echo-planar imaging factor = 55, sensitivity encoding factor = 2, field of view = 224 x 224 mm^2, voxel size = 2 x 2 x 2 mm^3 (matrix 112 x 112 pixels, 76 slices, transverse orientation). Additionally, six non-diffusion-weighted volumes were acquired to correct for echo-planar imaging-related geometrical distortions in anterior and posterior directions. All technical parameters of those six non-diffusion-weighted volumes were identical with those of the reference volume (non-diffusion-weighted) of the actual DTI sequence, except that for three volumes the readout direction was posterior-to-anterior and for the other three volumes the readout direction was anterior-to-posterior.

All preprocessing steps were fully automated using the diffusion MRI tools of the FMRIB Software Library (FSL; version 6.0, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki)

(Smith et al., 2004). First, the TOPUP tool (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/topup) was applied to the six non-diffusion-weighted volumes to correct for distortions of the susceptibility induced field. Second, the output from TOPUP was fed together with the DTI volumes into the EDDY tool (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy), which allows correction for subject's movements and eddy current-induced distortions. These preprocessed DTI data were then used for tract-based spatial statistics (TBSS) analyses.

Extraction of regions of interest

In order to restrict the statistical analyses of the TBSS data to perisylvian regions, we created a binary mask. Regions of interest (ROIs) were drawn from the Harvard-Oxford Cortical Structural Atlas (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases) as implemented in FSLVIEW (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslView). The following bilateral ROIs were selected based on the AP literature as well as on theoretical models of AP (Burkhard et al., 2019; Jäncke et al., 2012; Keenan et al., 2001; Loui et al., 2011, 2012; Schlaug et al., 1995; Schulze et al., 2009; Wengenroth et al., 2014): PT, PP, HG, as well as anterior and posterior divisions of the STG and MTG (Table 1). To obtain a binary mask that can be used for statistics with the RANDOMISE tool (see below), all 14 ROIs (unthresholded) were summed up.

Anatomic region	Hemisphere	MNI coordinates		
		X	**Y**	**Z**
Heschl's gyrus	L	-48	-18	5
	R	50	-16	6
Planum temporale	L	-54	-29	14
	R	56	-25	16
Planum polare	L	-48	-8	-4
	R	50	-6	-2
Anterior middle temporal gyrus	L	-58	-7	-20
	R	57	-4	-23

Posterior middle temporal gyrus	L	-59	-27	-13
	R	59	-22	-13
Anterior superior temporal gyrus	L	-56	-6	-9
	R	57	-3	-10
Posterior superior temporal gyrus	L	-57	-28	4
	R	59	-23	3

Table 3.1: Coordinates of the regions of interest (ROIs). The anatomic regions refer to the nearest-located cortical structures of the investigated white matter. For each ROI, the center of gravity is indicated with standard Montreal Neurological Institute coordinates (X, Y, Z). L = left hemisphere, R = right hemisphere.

Voxel-wise analysis of the fractional anisotropy data

Voxel-based analyses of the FA data were performed using TBSS (Smith et al., 2004), and the FDT toolbox (Behrens et al., 2003), which are both parts of the FSL library. To exclude non-brain tissue from further analyses, BET was applied to the non-diffusion-weighted images. Furthermore, the diffusion tensor model was fitted at each voxel using DTIFIT. Afterward, the TBSS steps were applied to the data. (1) FA data were prepared by using the –preproc option. (2) Nonlinear registration was applied to align the FA images to the high-resolution FMRIB58_BA standard-space image provided by FSL. (3) Using the –postreg option, the FA images of all subjects were affine-aligned to the 1 x 1 x 1mm MNI152 space. A skeleton was created by taking the mean of all FA images. (4) A threshold of 0.2 was applied to the mean FA skeleton image, which resulted in a binary mask. This binary mask contained the information of the voxels that were used in all successive processing steps. Finally, FA images of all subjects were projected onto the mean FA skeleton. This procedure resulted in an image file that contained the projected mean FA skeleton data (4D). The processed data were then subjected to non-parametric permutation statistics (RANDOMISE tool, see below).

Statistical analyses

For an overview of the surveyed behavioral measures, descriptive statistics are provided. Because no group classifications were conducted, means, standard deviations and variable ranges were calculated.

To reveal correlations between pitch-labeling scores and FA within AP related WM regions, the 4D skeletonized FA images were fed into voxel-wise statistics using RANDOMISE (Winkler et al., 2014) that applies nonparametric permutation testing. The null distribution was built upon 5,000 permutations. FA clusters that correlated with the pitch-labeling score were identified by the threshold-free cluster enhancement method. Furthermore, family-wise error (FWE) correction was applied to the data. To determine the labels of the brain areas that showed significant correlations with the pitch-labeling scores, the Harvard Oxford cortical structural atlas was applied. Only the labels with the highest probabilities are reported. For reasons of simplicity, we refer to the nearest-located cortical structures although the underlying WM was investigated. For visualization purposes, the fill command of TBSS was used to make the clusters graphically more visible.

For the voxel-wise TBSS analysis of the FA maps, effect sizes were reported. However, it must be noted that the computation is based on clusters that have already been identified as significant, resulting in an overestimation of the true effect (Kriegeskorte et al., 2010). All coordinates are reported in the standard Montreal Neurological Institute (MNI) space. To show the specificity of the results, a whole-brain analysis of the FA data is reported in the supplementary material (Table S1). The significance threshold for the analyses was set to $\alpha = 0.05$. Furthermore, to show results exceeding the main effects, the analysis was also conducted with a more liberate significance threshold ($\alpha = 0.20$, corrected for multiple comparisons).

3.5 Results

3.5.1 Behavioral data

Descriptive statistics of the variables age, musical training onset, estimated number of cumulative training hours across lifespan, musical aptitude, pitch-labeling score and cognitive capability are reported in Table 2. Furthermore, frequency distributions are given for the characteristics of sex and handedness.

Measures (N = 90)			
Sex (female/male)		45/45	
Handedness (right/left/ambidextrous)		79/7/4	
	Mean	Standard deviation	Range
Age (years)	26.8	4.78	18–39
Training onset[a] (years)	6.29	2.42	1–15
Musical training[b] (hours)	1.44	1.13	0.14–6.59
Musical aptitude[c]	65.1	6.42	45–77
Pitch-labeling score (%)	51.75	32.1	3.7–100
Cognitive capability[d]	126.62	31.53	20.83–198.68

[a] Age of musical training onset

[b] Estimation of the cumulative number of training hours across lifespan (units are given in 1 x 10^4)

[c] Total score of the AMMA-test

[d] Total score of the KAI-test

Table 3.2: Descriptive statistics of the demographic and behavioral variables. Mean, standard deviation, and range are depicted. For sex and handedness, the absolute frequencies are shown.

3.5.2 Results of the FA analyses

Voxel-wise analyses revealed a positive correlation between FA values and AP proficiency in a left-sided perisylvian cluster. This cluster (99 voxels in size, p = 0.024, r = 0.397, x/y/z = -39/-40/13, FWE corrected) comprised the WM underlying the left-sided PT (Figure 2). According to the JHU white-matter tractography atlas (Mori et al., 2005), the significant cluster corresponds to the posterior branch of the left-sided SLF.

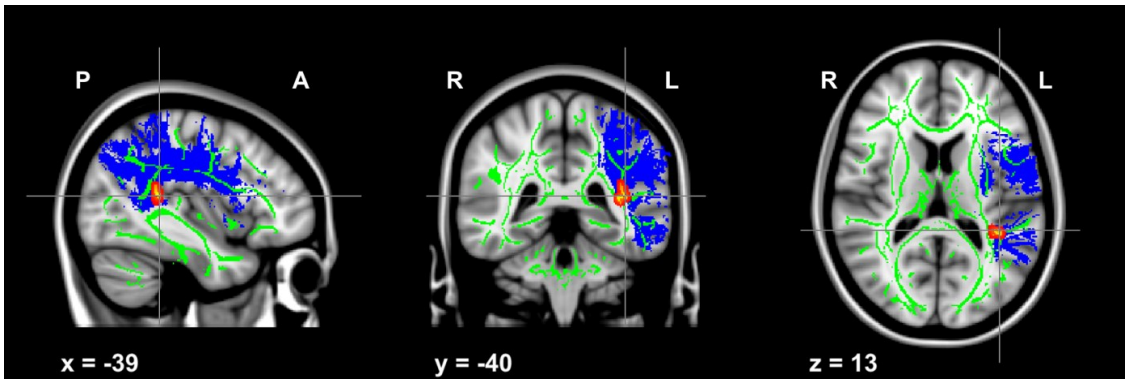

Figure 3.2: Results of the voxel-wise fractional anisotropy analyses. A cluster (red) located beneath the left planum temporale positively correlated with the pitch-labeling scores. The probabilistic tract structure of the left-sided superior longitudinal fasciculus is shown in blue. The superimposed crossline indicates the maximal values within the cluster as well as the corresponding Montreal Neurological Institute coordinates. Images are shown according to the radiological convention. A = anterior, P = posterior, L = left, R = right.

The correlation between the pitch-labeling scores and the mean FA values of the significant cluster of the WM underlying the left-sided PT (r = 0.397) is depicted in Figure 3.

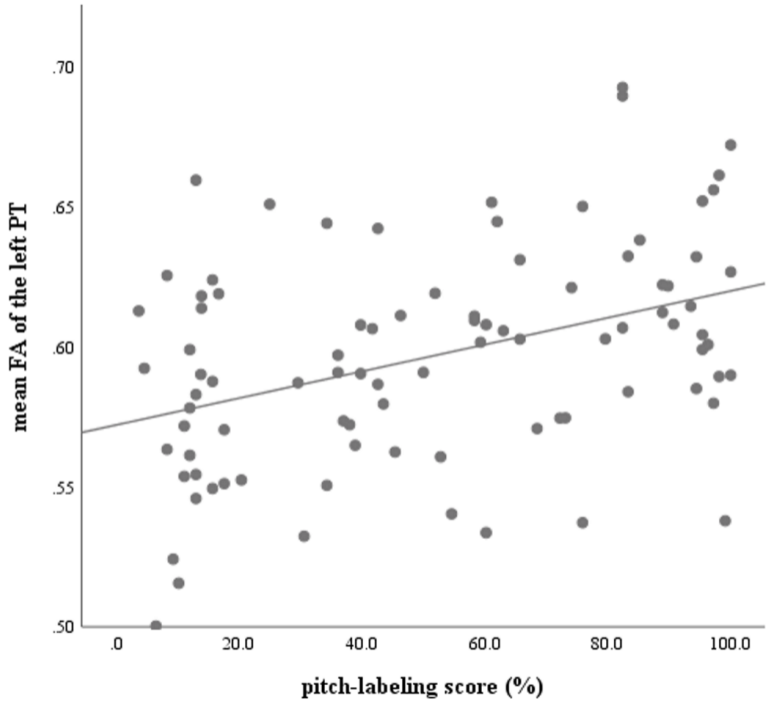

Figure 3.3: Correlation between the mean fractional anisotropy values of the significant cluster of the white matter underlying the left-sided planum temporale and the pitch-labeling scores (r = 0.397).

Furthermore, the analysis with a more liberal significance threshold of α = 0.20 (FWE corrected) revealed four additional smaller clusters, which are, except one cluster, located in the left hemisphere (Table 3). The largest of these additional non-significant clusters is located in the WM underlying the left PP (cluster 2). Furthermore, the analysis revealed a second cluster in the left-sided posterior PP at the border to the anterior HG (cluster 3). The only cluster located in the right hemisphere was also found in the WM underlying the PP (cluster 4). The smallest cluster contains eight voxels and was found in the insular cortex (cluster 5).

Cluster number	Cluster size (voxel)	p-value	MNI			Label[1]
			X	Y	Z	
1	301	.024	-39	-40	13	Planum temporale
2	75	.121	-43	-21	-8	Superior temporal gyrus / planum polare
3	34	.166	-37	-25	-1	Planum polare / insular cortex
4	10	.153	49	-11	-7	Planum polare
5	8	.198	-37	-18	-8	Insular cortex

[1] Labels were determined with the Harvard Oxford cortical structural atlas. The labels with the highest probabilities are depicted.

Table 3.3: Results of the voxel-wise analysis of the fractional anisotropy data with a significance threshold of $\alpha = 0.20$ (family-wise error corrected). As expected, the cluster size of the significant cluster increases with a more liberal threshold.

3.6 Discussion

The aim of the present study was to identify WM properties (here FA) of perisylvian brain areas that are associated with AP ability using a correlational approach. So far, only group differences between absolute and non-absolute listeners have been investigated. However, the intermediate levels of AP performance have almost been neglected. In this study, we examined whether possible gradual differences in pitch-labeling are also related to gradual differences in specific anatomical features of the WM. Based on previous studies, we hypothesized that FA would be higher in musicians with better pitch-labeling abilities. In line with this reasoning, we identified that FA values of the WM underlying the left PT positively correlated with AP ability.

Increased FA values are thought to reflect increased myelin integrity and WM organization (Schmithorst et al., 2002; Schmithorst & Wilke, 2002). Such an

optimization of microstructural WM properties could increase the velocity of impulse propagation between different patches of cortex and lead to a higher functional efficiency of information processing. With this in mind, one can argue that the more the PT and here particularly the left-sided part is efficiently connected to other brain regions within the auditory and non-auditory cortex, the better the AP ability is. Accordingly, previous studies have highlighted the importance of the PT for AP (Gaser & Schlaug, 2003; Keenan et al., 2001; Luders et al., 2004; Schlaug et al., 1995; Wilson et al., 2009). The PT has been proposed to be a "computational hub" for higher-order auditory processing. In order to be able to perform this function, the PT segregates the signals received from the primary auditory cortex into different spectro-temporal patterns using something similar to an independent component analysis (Griffiths & Warren, 2002; Hall & Plack, 2009). These signals are conveyed to higher brain areas for further processing. The spectro-temporal patterns relate to sound objects and their localization in space and are processed in the corresponding brain areas. Regarding the ability of AP, it could be that the PT allocates the different pitch chromas contained in the signal to different object patterns and forwards the information to higher-order auditory processing areas for further processing, which could enable the later occurring labeling of the different chromas. In addition, the PT is suggested to be involved in the processing of elementary pitch properties and pitch salience (Griffiths & Warren, 2002; Hall & Plack, 2009), which fits well with the possible role of the PT as a computational hub. Thus, it might be possible that the superior connection of this important hub subserves the establishment of AP ability.

As shown in Figure 2, the WM underlying the PT seems to be part of the SLF. The SLF is a fiber bundle that connects temporal with frontal brain regions and is especially involved in speech processing (Oechslin et al., 2009). Furthermore, the SLF plays a role in music processing. Bengtsson et al. (Bengtsson et al., 2005) found a positive correlation between FA values in the SLF and the estimated hours of music practice. In accordance with the dual-stream model of language perception and production (Rauschecker & Scott, 2009) the SLF is part of the postulated postero-dorsal stream.

Interestingly, the authors suggest that the dorsal processing stream originates from the PT or so-called pST (posterior superior temporal) region and that this brain area is involved in the spectro-temporal processing of acoustic sounds. Furthermore, Rauschecker and Scott (Rauschecker & Scott, 2009) pointed out that the postero-dorsal stream is also associated with "doable" templates, which representations are suggested to originate within the PT (Rauschecker & Scott, 2009; Warren et al., 2005). "Doable" should be interpreted in the context of the forward mapping mentioned in Rauschecker and Scott (2009). In particular, a heard sound is processed along the antero-ventral stream and then the information is transformed into a motor-action representation in the postero-dorsal stream, resulting in an efference copy for reproducing a specific sound. The postero-dorsal stream comes into play when a sound is produced and its production is monitored for the correct outcome. This interplay between forward and inverse mapping is not restricted to language. Halwani et al. (2011) found differences in the left dorsal pathways of singers, which they explained with the singers training requiring the forward and inverse mapping processes mentioned above. Hence, it could be that the ability of highly skilled APPs, e.g. to label and produce specific tones without reference, is modulated by the PT as part of the auditory postero-dorsal stream. However, increased FA values in the left-sided SLF in participants with higher AP abilities in the present study would confirm earlier findings from our group that showed a larger left-sided FA asymmetry in the SLF of APPs (Oechslin et al., 2009). Oechslin et al. (2009) found three clusters within the left-sided SLF that correlated with the performance in an AP test. One of the three clusters (x/y/z = -42/-50/4) lies in the vicinity of the cluster identified in the present study (x/y/z = -39/-40/13). The authors associated two of the clusters with the ability of APPs to relate a heard tone to their own instrument. Accordingly, the authors assumed that those two parts of the SLF underlie cortical regions that are important for the development of AP.

3.6.1 FUTURE PERSPECTIVES

In the present work, musicians with higher AP abilities revealed increased left hemispheric FA values in the PT. In general, the STG / superior temporal sulcus is thought to be involved in the long-term representation of pitch, pitch perception, categorization, and the identification of different auditory objects (Belin & Zatorre, 2000; Binder et al., 2000; Loui et al., 2012; Schulze et al., 2009; Wilson et al., 2009). However, previous studies showed a heterogeneous picture concerning the role of the STG in AP. For example, the anatomical study of Dohn et al. (2015) revealed an increased cortical thickness in the bilateral STG for APPs when compared to RPPs. In line with previous studies (Keenan et al., 2001; Luders et al., 2004; Schlaug et al., 1995), the authors reported a left lateralization of the pSTG for APPs. Furthermore, Loui et al. (2011) found enhanced FA for APPs in the bilateral WM pathways connecting the STG with the MTG. The same authors also reported a positive correlation between the left-sided tract volume and AP performance. A left-lateralization was also reported in the study by Oechslin et al. (2009), which showed increased FA in the left-sided SLF. Otherwise, Kim and Knösche (2016) revealed increased myelination in the right-sided PP in APPs. Furthermore, a positive correlation between the anterolateral part of the right-sided STG and a frequency discrimination threshold was found. A functional study by Schulze et al. (2009) showed a bilateral activation pattern in the STG for both, AP and non-AP musicians. A significant group difference was reported in the middle part of the left-sided superior temporal sulcus, and a trend towards significance for an increased bilateral activation in the STG for APPs. Increased small-world network parameters were found in the left-sided STG for APPs by Loui et al. (2012). A study by Wengenroth et al. (2014) highlighted the role of the right-sided perisylvian brain (including the dorsal STG) for APPs. Furthermore, a recently published EEG study from our lab (Burkhard et al., 2019) also revealed the relevance of the right perisylvian brain for APPs during attentive tone listening.

However, the studies mentioned above differed in terms of methods and paradigms used, which might explain a considerable amount of the heterogeneous results. The

findings mentioned above might highlight some aspects of a wider network that is required for AP processing involving both hemispheres: the left-sided brain could be presumably involved in labeling processes while bilateral or right-hemispheric perisylvian brain areas are recruited for perceptual aspects of AP processing, such as tone categorization. However, it is worth considering that the correlations found in this study relate more to labeling than to AP in general. This, due to the fact that the pitch-labeling test performance used for the calculations of the correlations relies on labeling a given tone correctly. In accordance, the left-hemispheric correlations between tract volume of the pathway connecting the STG with the MTG (Loui et al., 2011) and FA of the SLF (Oechslin et al., 2009) with AP performance can also been seen in that light. It could be that a left-sided perisylvian network might support the quick and effortless labeling process in APPs, which is, among others, reflected by increased FA. In accordance with this, Meyer and colleagues (2014) provided evidence for a relationship between increased cortical surface areas (CSA) and the corresponding WM of left-sided auditory related fields. The authors suggested that increased CSA and thinner cortex in auditory related areas could reflect the specialization of the left hemisphere for rapid temporal processing. Enhanced CSA could emerge from a wider spacing of microcolumns, which could lead to an advantage for the left hemisphere for processing temporal features of auditory stimuli (Hutsler, 2003). This view is reconcilable with previous models that suggested an advantage of left perisylvian brain areas for temporal processing (Poeppel, 2003; Zatorre et al., 2002). However, the authors (Meyer et al., 2014) suggested that thinner cortical thickness could be driven by enhanced myelination during brain development whereby gray matter is substituted by WM. From this perspective, the general basis of AP could be set during brain development.

However, future studies should incorporate both functional and anatomical methods and ideally also different paradigms (for example active and passive listening) to better elucidate the role of specific brain regions in different functions related to AP processing.

3.6.2 LIMITATIONS

A limitation of the present study could be that the pitch-labeling test was completed at home and not under strictly controlled laboratory conditions. It might be possible that the particular test procedure exerts an influence on the measured AP scores. However, unpublished studies in our lab revealed strong correlations between AP test scores obtained with online tests and under strict laboratory control. So far, there is no consensus about how to evaluate the AP ability reliably. Furthermore, previous studies differed with regard to the administration and scoring systems of the AP-tests that were used. Moreover, the stimulus material varied, for example, in type (sine tones, sampled instruments, etc.), length, loudness, and number of trials. As already mentioned in the introduction, there are various degrees of AP accuracy. Therefore, stimulus type, response accuracy and latency could provide valuable information to disentangle different labeling strategies. For example, already Bachem (1937) mentioned that some musicians possess a single intern reference tone (a specific tuning tone or the deepest tone one is able to sing) that is used to calculate the label of the target tone. This strategy would be reflected by longer reaction times as the computation needs more time than a labeling that is absolute and effortless. Furthermore, no standard scoring and analyzing method for the AP test evaluation is available. In some studies, only exact responses were counted as correct (Burkhard et al., 2019), whereas in other studies half-points were given for semi-tone deviations to the target tone (Psyche Loui et al., 2011). Both scoring strategies have consequences with respect to the distribution of test performance, as described by Bermudez and Zatorre (2009). However, it is not clear so far which would be the most valid technique to identify AP musicians. This will be an endeavor for future research. In conclusion, a progress in the field of AP research would be to reach a common agreement on the above-mentioned parameters in order to enable a more adequate and consistent evaluation of AP.

Based on the findings in the present study we suggest (as we and others did so previously) that the left PT and the fiber tracts connecting this hub with surrounding areas are related to AP. Furthermore, we provided evidence showing that FA values in the WM underlying the PT are gradually pronounced depending on AP skills.

The main aim of the present doctoral thesis is to explore AP ability in a relatively large sample. For this purpose, EEG and DTI data were collected from more than 100 participants to investigate functional as well as structural aspects of AP ability. Both studies presented here, lead to an improved understanding of the temporal and spatial neuronal characteristics of AP.

The aim of the EEG experiment (Study I) was to examine attentive listening in APP using either in tune or slightly mistuned tones as well as different noise segments. Participants were instructed to listen attentively to the presented tones. It was important for us to keep the listening experience as ecologically valid as possible. This was relevant because attentive listening to tones or melodies makes up the largest part of the daily musical routine of APPs. First, we investigated whether early or late auditory processing differences occur between APPs and RPPs, using the ERP technique. The temporal occurrence of the group differences in the ERPs is important to clarify the involved underlying neural mechanisms. More precisely, the time window contains information about the ongoing processes: while early stages may reflect pitch perception respective categorization, later stages could reflect more cognitive mechanisms such as pitch labeling or mnemonic processes. To investigate this issue, we analyzed several auditory evoked potentials, which served as biomarkers for the processes mentioned above. The investigated N100 and P200 components are thought to reflect early sensory and perceptual processes (Hillyard et al., 1998), while later endogenous and cognitive processes occur in a time window between 200–1500 ms after stimulus onset (Kutas & Federmeier, 2011) and were examined with the ERP components N200 and N400.

In the ERP analysis, we focused on the amplitudes of the single components, which were measured with the so-called signed area (Luck, 2014). Significant group

differences were only found for the early N100 component and a tendency toward significance for the other early component P200, in which APPs showed enhanced amplitudes compared to the RPPs. Both of the later components (N200 and N400) did not show significant group differences. In a second step, we wanted to know where in the brain those differences were generated. Therefore, a source estimation of the ERP component that showed significant group differences was calculated (N100). The source estimation revealed that the brain of the APPs shows enhanced right-sided perisylvian activation during the time frame of the N100 component, comprising the STS, MTG and HG. Results were interpreted as differences between APP and RPP at the stage of tone categorization, which is in line with previous findings (Hirata et al., 1999; Hirose et al., 2004; Schulze et al., 2009).

In a second experiment (Study II), DTI data of the same participants were investigated. Due to the results in Study I, we hypothesized that some perisylvian brain areas would show AP proficiency–dependent structural properties. Our particular focus was placed on FA, which is often interpreted as a marker of WM integrity or as an indicator of anatomical connectivity. We hypothesized that the musicians would show FA values of the WM underlying perisylvian brain structures depending on their scores in the pitch-labeling test. Furthermore, we expected the analysis to reveal a positive correlation in the sense that a higher score would lead to higher FA values. A significant positive correlation was found in the left-sided PT. This gives further emphasis to the previously shown importance of this structure for AP (Gaser & Schlaug, 2003; Keenan et al., 2001; Luders et al., 2004; Wilson et al., 2009). The finding in the left-sided PT presumably reflects aspects of categorization processes, which are mandatory for the subsequent labeling process.

Whether and how these results further extend our knowledge of AP ability is discussed below. Furthermore, methodological considerations with regard to the pitch-labeling test and the methods used in the studies presented in this thesis are described. Finally, attempts for improvements and a conclusion are provided.

4.1 AP PSYCHOLOGICAL ASPECTS

In the following sections, the knowledge gained with the studies presented in this doctoral thesis is summarized regarding AP ability. First, the phenomenon of categorization is discussed and how it could be related to AP ability. Afterward, a short section deals with the labeling process. Finally, we discuss how AP could relate to attention.

4.1.1 ABSOLUTE PITCH AND CATEGORIZATION AND LABELING

Within the scope of AP, categorization should be interpreted in the sense that an auditory signal is divided into specific, well-defined units that correspond to the pitch chromas (Bachem, 1955; Miyazaki, 1989; Miyazaki, 1988; Takeuchi & Hulse, 1993). This means that the sounds can be perceived as something like auditory "individuals", which have their own characteristics. In this way, an APP is able to distinguish specific pitches without the need of a reference tone. Furthermore, it can be assumed that the categorization of APPs primarily enables the identification of the chroma of a sound: that is, which acoustic characteristic makes a "G" recognizable as such. Indeed, it appears that the information relating to the height of a pitch is less paramount because APPs cannot always determine the correct octave of a correctly recognized tone (Miyazaki, 1989). Together, both studies presented in this thesis provide further evidence for an AP-specific categorization process relevant to APPs if they listen attentively to a sound. Dittinger et al. (2016) associated the N200 component with the auditory categorization process and not the earlier occurring N100 component that significantly differed between the AP and RP groups in Study I. However, the conducted source estimation for the N100 component revealed that APPs showed enhanced activity in two brain areas that are closely related to auditory categorization (STS and MTG), which points to a very early occurring AP-specific categorization process. Klein and Zatorre (2011) have already associated STS activation with the categorization of musical chords in non-APPs. However, it is conceivable that this brain

area is more specialized in APPs and in addition to the categorization of the sound pattern of a musical chord, is moreover involved in the categorization of a single pitch chroma. This assumption is reinforced by the enhanced activation of the MTG in APPs compared to RPPs in Study I. As already suggested by Loui et al. (2011, 2012), the MTG is able to access already-stored templates of pitch categories. Taken together, the very early activation of these two areas around 100ms after stimulus onset could enable a very fast and efficient pitch categorization in APPs. Furthermore, the enhanced FA values in the PT in participants with a high pitch-labeling score in Study II could also be parsimoniously interpreted in the sense of an AP-specific categorization process. Based on the postulated function of the PT as a computational hub (Griffiths & Warren, 2002; Hall & Plack, 2009), we assume that the PT endows the preprocessed information from the early stages of the auditory processing stream with object patterns and forwards it to higher brain areas. In this way, pitch chroma information is transformed into distinguishable objects, which can be labeled in a later processing step. The positive correlation between AP proficiency and FA values found in Study II could support the assumption that high AP proficiency requires efficient use of this hub function of the PT. Furthermore, high FA could possibly be understood in terms of an optimized WM structure (Schmithorst et al., 2002; Schmithorst & Wilke, 2002), which enables a faster and more efficient connection to other brain areas. Since many studies that have investigated passive listening in APPs did not find any evidence for a categorization process (Elmer et al., 2013; Pantev et al., 1998; Wu et al., 2008), we assume that attention could be the decisive factor. Furthermore, it could be possible that this categorization process involves bilateral perisylvian brain structures since our studies revealed the involvement of the left (structural Study II), as well as the right hemisphere (functional Study I). The attentive perception of a tone could be the prerequisite for the initiation of the categorization process and thus enable the subsequent labeling. However, it seems that the categorization process is not automatically linked to the labeling. Study I was designed in such a way that the participants did not have to solve a specific task, and the results do not provide any clues for labeling to occur. In sum, the early categorization process could thus be the crucial component for AP, on which the

later and optional labeling could depend. As already mentioned, different labeling strategies in RPPs could also lead to high pitch-labeling performance (Bachem, 1937) and by definition, labeling becomes an AP typical characteristic only, when it is carried out automatically and without any effort (Levitin & Rogers, 2005). The results of Study I suggest that the labeling does not necessarily occur automatically. Moreover, it seems that the labeling is closely linked to the categorization process, which commences much earlier. Moreover, Study I clearly shows that the labeling need not occur for differences between RPPs and APPs to be demonstrated. The labeling appears to be an independent module that the APPs can call if necessary. However, the enhanced FA values in Study II could also be interpreted as part of a labeling process because we calculated the correlation with the pitch-labeling score. If so, the results of Study II extend those of Study I in a sense that the increased AP proficiency, which is at least partially made possible by the categorization process, could lead to a better-connected PT. In this way, the PT, with its function as a computational hub (Griffiths & Warren, 2002; Hall & Plack, 2009), could enable the forwarding of the already categorized information to those brain areas that are involved in the labeling process.

4.1.2 Absolute pitch and attention

The results in Study I imply that AP at least partially depends on attention. This conclusion is not only based on the results of Study I but also on previous studies that revealed differences in early ERP components (N100/P200) during attentive listening. Moreover, this view is supported by studies that did not find any differences in early ERP components during passive listening (Elmer et al., 2013; Pantev et al., 1998; Wu et al., 2008). Therefore, we assume that the categorization process in APPs is automatically initiated, but only if the APP drives its attention toward the auditory stimuli. However, Schulze et al. (2009) suggest that AP ability comprises two important components: an early categorization and a later labeling component. In contrast, we suggest that the AP-specific labeling component itself is comprised of two different processes. First, we assume that an automatically running categorization process occurs

early in the auditory processing stream, as also suggested by Schulze et al. (2009). We hypothesize that this categorization process is only initiated if APPs turn their attention to acoustic stimuli. Furthermore, we assume that the categorization process is a necessary precondition for the second process. The second process consists of forwarding the already categorized acoustic information to enable the labeling. The quality of the AP-specific labeling could depend on the quality of the preceding categorization process. Furthermore, the labeling itself does not occur automatically but rather is initiated only if the labeling of single tones is needed. If so, it could be that the AP-specific labeling may differ from other labeling strategies through the dependency on the categorization process, wherein other labeling strategies can be accomplished, for example, through the use of a single internal reference tone (Bachem, 1937). In summary, the execution of AP-specific tasks (labeling or producing a tone without reference) may require that the APP be listening attentively. This attentive listening in turn automatically induces the categorization process, which enables the labeling when required. From that perspective, AP ability resembles the seeing and labeling of colors: most of us perceive different colors but we are normally not constantly engaged in labeling them. Clearly, such labeling would unnecessarily consume too much of our substantial resources. Only if we are asked, for example, what color a passing car has, do we label the visual impression.

4.2 Methodological considerations

The following section provides a critical evaluation of the methods used in Studies I and II. First, some reflections regarding the pitch-labeling test are described. Afterward, we delineate the pros and cons of the methods used, especially the signed area ERP technique, which is a newer technique and therefore has not frequently been applied so far. This section closes with a short attempt to highlight the discrepancies and correspondences of the functional and structural methods used.

4.2.1 ABSOLUTE PITCH CLASSIFICATION

As already mentioned in the introduction of this thesis, the evaluation of AP ability is one of the main challenges concerning AP research. Parts of the heterogeneous findings in AP research could be due to the fact that up to now there has been no standard procedure for classifying AP ability. A wide variety of factors would have to be standardized; for example, the awarding of points, the stimulus choice, and many other influential parameters. This point was raised by Bermudez and Zatorre in 2009. Unfortunately, the research community is far from developing a standard method.

According to Bermudez and Zatorre (2009), a mandatory parameter for classifying AP is the elicitation of unbiased and precise reaction time measurements, whereby a precise reaction time measurement is inter alia highly dependent on the latency and bandwidth of the internet connection and therefore inevitably lead to performance differences. Therefore, it is strongly recommended to perform the AP test under strictly controlled laboratory conditions (ideally using an offline method) if the classification of AP is verified explicitly and does not rely on self-report of the participants. Furthermore, reaction time measurements could be helpful to uncover other, more time-consuming labeling strategies than the absolute judgment of pitch. For example, calculating the target tone in relation to a single or a few available internal reference tones is a strategy that costs time. For this reason, reaction time measurements could potentially be used to identify different subtypes of AP that are not distinguishable on the sole basis of the pitch-labeling score. Another possibility, at least to distinguish absolute from (at least partially) relative labeling strategies, could be to introduce a certain time limit for the answer. However, this would be disadvantageous in two ways. First, it is not unambiguous to set the temporal cut-off. How much time an APP needs to label a heard tone that clearly allows a dichotomy classification of the labeling strategy is poorly defined. Second, with this strategy, one would produce missing values (as was the case in the pitch-labeling test conducted in the project presented here), which could be potentially unfavorable as outlined below. Introducing a forced-choice test instead could at least solve the latter problem. Furthermore, the use of stimuli that are presumably

equally familiar, or even better, equally non-familiar to all participants should be favored, to avoid instrumental cues that could be used by only those who are familiar with the given stimuli. However, a further critical factor lies in the applied scoring strategy (Bermudez & Zatorre, 2009). The most widely used method (also used in the studies presented in this thesis) is to exclusively assign a point for the exact identification of the presented target tone. A further common scoring strategy is to award full or fractional points for answers that lie in the range of semitone deviations to the target tone. These arbitrary choices might seem innocuous at first sight, but are able to affect some nontrivial aspects: high performers, whose answers are consistently shifted (e.g. due to paracusis: recent listening experiences that temporally change the internal tuning, or individual learning history) would be perceived as performing relatively poorly. Furthermore, according to Bermudez and Zatorre (2009), counting only exactly correct answers (and thereof neglecting the variance) could lead to an overestimation of the sometimes described bimodal distribution (Athos et al., 2007; Carroll, 1975). The authors argue that the bimodal impression might be exaggerated as a consequence of this scoring strategy by its artificial sharpening of the borders of the supposed distributions. On the other hand, the image is blurred if fractional points are assigned, preventing the distinguishing of perfect performers from those centering around the target tone with some deviations. For these reasons, among others, the authors suggest using the so-called mean absolute deviation (MAD) score in combination with reaction time measurements, particularly to identify intermediate performers. As the name implies, the MAD is the mean of the absolute value of the deviation from the target tone, in semitone steps. Because the octave information is neglected, the MAD score ranges from -6 to +6, where a MAD score of 3 represents a random response pattern and a MAD of 0 reflects a perfect performance without any deviations. To calculate the MAD index, it is necessary to implement a forced-choice method because (a) one cannot assign a score that simply describes a wrong answer, and (b) the possibility of missing values penalizes those who answer every single trial even if they are not sure about the correct answers, while participants only answering those trials of which they are sure will receive a higher score. This holds true even if only a

few trials are answered. This is the reason why we refrained from calculating the MAD in the here-presented Study I: unfortunately, our pitch-labeling score compromised missing values. However, although not a perfect solution, classification based on self-report (as was done in Study I) often corresponds to several indices, as shown by Bermudez and Zatorre (2009). Furthermore, one of the greatest strengths of this classification technique lies in the fact that an arbitrarily chosen cutoff can be avoided. Moreover, the recruitment explicitly addressed professional musicians regardless of AP ability and there were no obvious benefits for the participants whether they showed AP ability or not.

4.2.2 Event-related potential technique: signed area

As already described in the introduction, the ERP technique used in Study I (signed area) is a relatively new approach, and only a few studies have used this technique until now. Therefore, it seems reasonable to discuss the advantages and disadvantages of the different measurements of amplitudes that already exist and are widely used (peak and mean amplitude). In a second step, we compare the signed area measure to the conventional techniques and highlight why we considered signed area to be the most suitable measure for Study I. To summarize, peak amplitude is defined as the maximum voltage point, mean amplitude takes the average voltage over a predefined time window, and signed area is the area above or below the baseline.

First, it is important to keep in mind that an ERP component extends through time. Accordingly, an important consideration to take into account when interpreting ERP waveforms was mentioned by Luck (Luck, 2014):

"Peaks and components are not the same thing. There is nothing special about the point at which the voltage reaches a local maximum." (Luck, 2014, p. 286)

Seen in this light, mean amplitude is a better alternative than peak measurements. Another major advantage over peak amplitude measurements is that mean amplitude is relatively insensitive to high-frequency noise. This is because small noise peaks cancel each other out if a fairly large time window is defined. To overcome this weakness in peak measurements, a time window can be defined around the peak, and an average voltage around the peak could be calculated. However, this does not solve the problem of defining a reliable point as the peak amplitude. Furthermore, even if a high-cut off filter is applied to the data, it is common that visually finding the peak of a specific component is not always unambiguous. As shown in Figure 1.2, identifying the peak of the P3 component is easy for Subject 1 but it is almost impossible to find a representative peak in Subjects 2 and 3. If we were to rigidly follow the rules for identifying the local peak amplitude for each subject, we would run into another difficulty: as indicated by the arrows in Figure 1.2, the time points for the P3 peak amplitude would differ greatly between the subjects (Subject 1 ≈ 500 ms; Subject 3 ≈ 700 ms). This represents a problem because it is unlikely that the neurophysiological processes that induce a P3 component differ in timing by approximately 200 ms, at least in healthy subjects. The question that arises is whether we still measure the same signal or rather merely the same processes. The same holds true not only for different subjects but also when comparing different conditions or different electrode sites at different time points. Hence, mean amplitude is clearly superior in the respect that with mean amplitude the same time points for all subjects, conditions, and/or electrodes enter the computation. However, in cases where latency differences occur, one should consider trying difference waves whenever possible to isolate single components or, as a second option, using signed area, which is described below. However, when comparing groups or conditions with different amounts of noise in the data, we should also tend to measure mean amplitude or signed area rather than peak amplitudes. One factor influencing the amount of noise in the data is the number of trials. As mentioned before, the more trials enter the analysis, the better the signal-to-noise ratio. Hence, the trial count has an influence on the noise level, which means we must take care when comparing groups or conditions with different numbers of trials. Of course, other

sources of noise do exist as well: for example, comparing specific patient groups to healthy controls or when comparing children to adults, and so on. However, with increasing noise level, the maximum amplitude, on which peak measures are based, will also increase. Mean amplitude, on the other hand, does not become larger the noisier the data are, although noise could lead to decreased statistical power due to more variability in the data.

The main advantage of using signed area is that we do not rely on narrow time windows. In other words, it does not matter in most cases which time window is defined, as long as the window is wide enough. This procedure prevents the researcher from choosing an arbitrary time window that may introduce a bias toward significant results. Furthermore, the ERP component is treated as extending through time, without superposition (and resulting cancelation) by neighboring components (as would be the case for mean amplitude). In most cases, it is possible to fully capture the area of a specific component and not only a part thereof. Due to the large time window used, signed area is also able to catch activity that occurs at different time points between groups, conditions, or electrode sites. The downside of using signed area is that noise deflections also enter the computation. As is the case for peak amplitude, noise influences area measures in the way that the area values increase. Furthermore, the range of the area values is not intuitively accessible, as it is for peak or mean amplitude. In other words, area values no longer correspond to the initial unit and range of the raw amplitudes. Moreover, the usage of area measures is relatively recent and is not yet implemented in popular software packages such as BrainVision Analyzer. As a consequence, little support is available and the researcher therefore has to think carefully about the method and what to do with it.

4.2.3 Functional and structural conformities

While Study I revealed right-hemispheric findings, Study II showed the opposite pattern: respective hemispheric lateralization. However, it should be considered that in

the functional Study I, no task was implemented, while the results in the structural Study II depend on the pitch-labeling abilities derived from the pitch-labeling task. Seen this way, the results of Study I may reflect more perceptual aspects of AP, while the results of Study II are particularly based on pitch labeling. Moreover, it has to be considered that the relation between brain function and structure is not perfectly clear. Hence, care must be taken not to over-interpret the correlation between the FA values in the left PT and the AP score revealed by Study II. Nonetheless, brain function can unequivocally be influenced by the underlying structural brain architecture (Hou et al., 2017; Liem et al., 2012). For example, Chu et al. (2015) investigated inter alia the concordance of resting-state EEG networks and DTI networks. The authors showed several correspondences between those networks. The study revealed that long-distance networks between brain areas in the gamma range corresponded to underlying structural WM networks. Furthermore, no such correspondences were found for lower frequency bands. The authors assume that amplitudes in higher frequencies (such as gamma oscillations) enable the functional coupling or binding of more distal brain areas if and only if corresponding WM paths between those brain areas do exist. However, because of its role as a computational hub, it is likely that the PT also comprises long-distance connections to other brain areas, such as frontal areas, which are addressed, for example, to accomplish the labeling process. Zatorre et al. (2002) indicate in their review study that according to a previous postmortem study by Anderson et al. (1999), the left-sided PT showed more WM than the right-sided PT. According to Hutsler and Gazzaniga (1996), this asymmetry occurs together with larger pyramidal cells in the left primary auditory cortex. The cell columns formed from these pyramid cells are also larger and more widely spaced (Seldon, 1981a, 1981b, 1982). Furthermore, the authors suggested that these cells in the left auditory cortex are in general better connected than those on the right side. This, in turn, leads to more efficient and especially faster signal transmission in the left perisylvian brain and thus increases the temporal resolution, which is functionally manifested in better processing of speech discrimination in the left hemisphere (Zatorre et al., 2002). High temporal resolution signals would roughly correspond to signals in the gamma oscillations, which were thought to play a key role

in the study by Chu et al. (2015). Therefore, according to Zatorre et al. (2002), it is likely that information processing in the left PT occurs in the gamma range. Taken together, the higher FA values in the PT with the increasing pitch-labeling proficiency revealed in Study II could be an indication of long-distance structural connections to other brain areas, in this case probably to frontal ones.

In addition, there is evidence that AP-specific spectral processing steps could occur in the left hemisphere. PET studies with speakers of tonal languages were able to show that pitch contour processing and speech discrimination occurs in the left hemisphere (Gandour et al., 1998, 2000; Klein et al., 2001).

4.3 Future perspectives

Based on the knowledge gained in the present doctoral thesis, future studies should inter alia try to isolate the AP categorization template. For example, the template could be addressed as follows: one could manipulate everyday noises (e.g. a door being slammed) in a way that the noise is composed of a single pure tone instead of the common interplay of different pitches. The idea behind this is that those noises would trigger the template in APPs and the sound would thus be categorized. With RP, on the other hand, it is very unlikely that their strategies, which they otherwise use when listening to sound, will also come into play here since everyday noises are no longer in their area of expertise (music). Furthermore, two additional experimental conditions could help to isolate the AP template: first, a natural, non-manipulated everyday noise condition and second a tone listening condition similar to the attentive listening condition presented in Study I. For each condition and group, ERPs would be calculated. Afterward, one could calculate difference waves by subtracting the tone listening condition from the manipulated everyday noise condition. One possible hypothesis could be that the amplitude of the difference wave would be higher in RPP because for these conditions they would use completely different listening strategies. For the APPs, on the other hand, one would expect that the amplitudes of the ERPs of

both conditions would not show large differences, at least within the first 200 ms of stimulus onset, because the stimuli of both conditions would trigger their categorization template. Furthermore, if one were to calculate difference waves by subtracting the manipulated from the natural everyday noise conditions, one could potentially isolate the AP template because the categorization template would only be activated for the manipulated everyday noise but not for the natural everyday noise stimuli. For the RPPs, these two conditions should not show any large amplitude differences because the stimuli of both conditions would consist solely of everyday noises and therefore no expert knowledge would be demanded.

Furthermore, as already mentioned above, future research must focus on the development of a standard method to evaluate AP ability. In a second step, it must to be clarified whether AP: (a) represents a distinct population within the population of musicians, (b) lies on one end of a continuous ability of tone identification, or (c) whether APPs indeed form a distinct population within the population of the musicians such that the intermediate performance levels could be explained with a continuous performance level within this population of APPs. Although it is difficult to recruit a large number of APPs by reason of low incidence, this is precisely what future studies should seek to do. As already discussed, the more participants involved, the higher the statistical power. High statistical power is able to enhance the confidence that the results veridically represent the underlying population of APPs. Another way to gain confidence in research results is to perform a large number of replication studies. The execution of replication studies is challenging for several reasons but is unavoidable in the research domain. Due to the Koole and Lakens (2012), only in the long-run are we able to estimate whether a given result represents a true effect or not. This latter point is not as trivial as it may first appear: researchers aim to publish new insights and are therefore always looking for novel effects. Replication studies, in turn, do not guarantee a promising and glorious career as a researcher. However, collaborations and data exchange should be the gold standard in future research, alongside pre-registration, to increase the credibility of the results.

In summary, the present doctoral thesis provides further evidence that the phenomenon of AP is highly complex. Study I revealed that AP-specific processing steps occur very early within the auditory processing stream, specifically during the first 100 ms of acoustic stimulus onset. Differences in the amplitude of the N100 ERP component between APPs and RPPs imply differences in the early categorization process. Furthermore, the increase in FA values in the PT with increasing pitch-labeling proficiency shows the importance of this brain area for AP ability, as already highlighted by many previous studies. Study I and Study II revealed results that speak to different hemispheric lateralization. This could be related to the assumption that AP is more complex than hitherto assumed, whereby different methods and tasks or conditions are able to only highlight particular aspects instead of the overall picture of this ability. For this reason, the goal of future studies investigating AP should be, apart from large sample sizes, the conducting of different experiments (as was done in the project that this thesis is a part of) and the replication of those in different laboratories around the world. AP seems to be more than just the sum of some single processes. Rather, it is a complex phenomenon that is embedded in several cultural, psychological, and physiological aspects of its possessors. To unravel its true nature, it has to be treated as such a complex skill.

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*, *4*(3), 316–329.

Anderson, B., Southern, B. D., & Powers, R. E. (1999). Anatomic asymmetries of the posterior superior temporal lobes: a postmortem study. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *12*(4), 247–254.

Annett, M. (1970). A classification of hand preference by association analysis. *British Journal of Psychology* , *61*(3), 303–321.

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschier, J. (2007). Dichotomy and perceptual distortions in absolute pitch ability. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(37), 14795–14800.

Bachem, A. (1937). Various Types of Absolute Pitch. *The Journal of the Acoustical Society of America*, *9*(2), 146–151.

Bachem, A. (1940). The Genesis of Absolute Pitch. In *The Journal of the Acoustical Society of America* (Vol. 11, Issue 4, pp. 434–439). https://doi.org/10.1121/1.1916056

Bachem, A. (1955). Absolute Pitch. *The Journal of the Acoustical Society of America*, *27*(6), 1180–1185.

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute pitch: an approach for identification of genetic and nongenetic components. *American Journal of Human Genetics*, *62*(2), 224–231.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. In *Biophysical Journal* (Vol. 66, Issue 1, pp. 259–267). https://doi.org/10.1016/s0006-3495(94)80775-1

Baumann, S., Meyer, M., & Jäncke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *Journal of Cognitive Neuroscience*, *20*(12), 2238–2249.

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., Matthews, P. M., Brady, J. M., & Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *50*(5), 1077–1088.

Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A., & Larson, C. R. (2014). Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain and Cognition*, *84*(1), 97–108.

Belin, P., & Zatorre, R. J. (2000). "What", "where" and "how" in auditory cortex. *Nature Neuroscience*, *3*(10), 965–966.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, *8*(9), 1148–1150.

Bennett, E. L., Diamond, M. C., Krech, D., & Rosenzweig, M. R. (1996). Chemical and anatomical plasticity of brain. 1964. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *8*(4), 459–470.

Bennett, M. V. L. (1971). Electric Organs. In *Fish Physiology* (pp. 347–491).

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex* , *19*(7), 1583–1596.

Bermudez, P., & Zatorre, R. J. (2005). Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *1060*, 395–399.

Bermudez, P., & Zatorre, R. J. (2009). A Distribution of Absolute Pitch Ability as Revealed by Computerized Testing. *Music Perception*, *27*(2), 89–101.

Bermudez, P., & Zatorre, R. J. (2009). The absolute pitch mind continues to reveal itself. *Journal of Biology*, *8*(8), 75.

Besson, M., & Kutas, M. (1993). The many facets of repetition: a cued-recall and event-related potential analysis of repeating words in same versus different sentence contexts. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *19*(5), 1115–1133.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex* , *10*(5), 512–528.

Broca, P. (1861). Remarques sur le siége de la faculté du language articulé, suivies d’une observation d'aphémie (perte de la parole). *Bulletin de La Société Anatomique*, *6*, 330–357.

Bryden, M. P. (1977). Measuring handedness with questionnaires. *Neuropsychologia*, *15*(4-5), 617–624.

Burkhard, A., Elmer, S., & Jäncke, L. (2019). Early tone categorization in absolute pitch musicians is subserved by the right-sided perisylvian brain. *Scientific Reports*, *9*(1), 1419.

Carroll, J. B. (1975). Speed and accuracy of absolute pitch judgments: some latter-day results. In *ETS Research Bulletin Series* (Vol. 1975, Issue 2, pp. i – 71). https://doi.org/10.1002/j.2333-8504.1975.tb01075.x

Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. In *Annals of Neurology* (Vol. 57, Issue 1, pp. 8–16). https://doi.org/10.1002/ana.20319

Chu, C. J., Tanaka, N., Diaz, J., Edlow, B. L., Wu, O., Hämäläinen, M., Stufflebeam, S., Cash, S. S., & Kramer, M. A. (2015). EEG functional connectivity is partially predicted by underlying white matter connectivity. *NeuroImage*, *108*, 23–33.

Dan, Y., & Poo, M.-M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, *44*(1), 23–30.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980.

Deutsch, D. (2002). The Puzzle of Absolute Pitch. In *Current Directions in Psychological Science* (Vol. 11, Issue 6, pp. 200–204). https://doi.org/10.1111/1467-8721.00200

Deutsch, D., Dooley, K., Henthorn, T., & Head, B. (2009). Absolute pitch among students in an American music conservatory: association with tone language fluency. *The Journal of the Acoustical Society of America*, *125*(4), 2398–2403.

Deutsch, D., Henthorn, T., & Dolson, M. (2004). Absolute Pitch, Speech, and Tone Language: Some Experiments and a Proposed Framework. *Music Perception*, *21*(3), 339–356.

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. *The Journal of the Acoustical Society of America*, *119*(2), 719–722.

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional Music Training and Novel Word Learning: From Faster Semantic Encoding to Longer-lasting Word Representations. *Journal of Cognitive Neuroscience*, *28*(10), 1584–1602.

Dohn, A., Garza-Villarreal, E. A., Chakravarty, M. M., Hansen, M., Lerch, J. P., & Vuust, P. (2015). Gray- and white-matter anatomy of absolute pitch possessors. *Cerebral Cortex* , *25*(5), 1379–1388.

Elmer, S., Rogenmoser, L., Kühnis, J., & Jäncke, L. (2015). Bridging the gap between perceptual and cognitive perspectives on absolute pitch. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(1), 366–371.

Elmer, S., Sollberger, S., Meyer, M., & Jäncke, L. (2013). An Empirical Reevaluation of Absolute Pitch: Behavioral and Electrophysiological Measurements. *Journal of Cognitive Neuroscience*, *25*(10), 1736–1753.

Ericsson, K. A., Anders Ericsson, K., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. In *Psychological Review* (Vol. 100, Issue 3, pp. 363–406). https://doi.org/10.1037/0033-295x.100.3.363

Fields, R. D. (2005). Myelination: an overlooked mechanism of synaptic plasticity? *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *11*(6), 528–531.

Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., & Sandmann, P. (2016). On the relationship between auditory cognition and speech intelligibility in cochlear implant users: An ERP study. *Neuropsychologia*, *87*, 169–181.

Gaab, N., Schulze, K., Ozdemir, E., & Schlaug, G. (2006). Neural correlates of absolute pitch differ between blind and sighted musicians. *Neuroreport*, *17*(18), 1853–1857.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, *12*(1), 207–222.

Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport*, *9*(9), 2115–2119.

Gaser, C., & Schlaug, G. (2003). Gray matter differences between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *999*, 514–517.

Gazzaniga. (2013). *Cognitive Neuroscience the Biology of the Mind*.

Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, *56*(6), 1127–1134.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Gordon, E. E. (1989). *Manual for the Advanced measures of music audiation*.

Gramfort, A., Papadopoulo, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomedical Engineering Online*, *9*, 45.

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (1999). Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component [Review of *Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component*]. *American Journal of Human Genetics*, *65*(3), 911–913.

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (2001). Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment. *American Journal of Medical Genetics*, *98*(3), 280–282.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*(7), 348–353.

Hall, D. A., & Plack, C. J. (2009). Pitch processing sites in the human auditory brain. *Cerebral Cortex* , *19*(3), 576–585.

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of Practice and Experience on the Arcuate Fasciculus: Comparing Singers, Instrumentalists, and Non-

Musicians. In *Frontiers in Psychology* (Vol. 2). https://doi.org/10.3389/fpsyg.2011.00156

Hänggi, J., Beeli, G., Oechslin, M. S., & Jäncke, L. (2008). The multiple synaesthete E.S.: neuroanatomical basis of interval-taste and tone-colour synaesthesia. *NeuroImage*, *43*(2), 192–203.

Hebb, D. O. (1949). *The Organisation of Behaviour: A Neuropsychological Theory*.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*(5), 393–402.

Hillyard, S. A., Teder-Sälejärvi, W. A., & Münte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, *8*(2), 202–210.

Hirata, Y., Kuriki, S., & Pantev, C. (1999). Musicians with absolute pitch show distinct neural activities in the auditory cortex. *Neuroreport*, *10*(5), 999–1002.

Hirose, H., Kubota, M., Kimura, I., Ohsawa, M., Yumoto, M., & Sakakihara, Y. (2002). People with absolute pitch process tones with producing P300. *Neuroscience Letters*, *330*(3), 247–250.

Hirose, H., Kubota, M., Kimura, I., Yumoto, M., & Sakakihara, Y. (2003). N100m in children possessing absolute pitch. *Neuroreport*, *14*(6), 899–903.

Hirose, H., Kubota, M., Kimura, I., Yumoto, M., & Sakakihara, Y. (2004). N100m in adults possessing absolute pitch. *Neuroreport*, *15*(9), 1383–1386.

Hirose, H., Kubota, M., Kimura, I., Yumoto, M., & Sakakihara, Y. (2005). Increased right auditory cortex activity in absolute pitch possessors. *Neuroreport*, *16*(16), 1775–1779.

Hou, J., Jiancheng, H. O. U., Song, B., Zhou, J., Changan, S. U. N., & Haidong, Z. H. U. (2017). The cognitive and neural mechanisms of absolute pitch. *Advances in Psychological Science*, *25*(11), 1865.

Hutsler, J. J. (2003). The specialized structure of human language cortex: pyramidal cell size asymmetries within auditory and language-associated regions of the temporal lobes. *Brain and Language*, *86*(2), 226–242.

Hutsler, J. J., & Gazzaniga, M. S. (1996). Acetylcholinesterase staining in human auditory and language cortices: regional variation of structural features. *Cerebral Cortex* , *6*(2), 260–270.

Itoh, K., Suwazono, S., Arao, H., Miyazaki, K. 'ichi, & Nakada, T. (2005). Electrophysiological correlates of absolute pitch and relative pitch. *Cerebral Cortex* , *15*(6), 760–769.

Jäncke, L. (2005). *Methoden der Bildgebung in der Psychologie und den kognitiven Neurowissenschaften*. Kohlhammer Verlag.

Jäncke, L. (2017). *Lehrbuch Kognitive Neurowissenschaften*.

Jancke, L., Baumann, S., Koeneke, S., Meyer, M., Laeng, B., Peters, M., & Lutz, K. (2006). Neural control of playing a reversed piano: empirical evidence for an unusual cortical organization of musical functions. *Neuroreport*, *17*(4), 447–451.

Jancke, L., Buchanan, T. W., Lutz, K., & Shah, N. J. (2001). Focused and nonfocused attention in verbal and emotional dichotic listening: an FMRI study. *Brain and Language*, *78*(3), 349–363.

Jäncke, L., Gaab, N., Wüstenberg, T., Scheich, H., & Heinze, H. J. (2001). Short-term functional plasticity in the human auditory cortex: an fMRI study. *Brain Research. Cognitive Brain Research*, *12*(3), 479–485.

Jäncke, L., Langer, N., & Hänggi, J. (2012). Diminished whole-brain but enhanced peri-sylvian connectivity in absolute pitch musicians. *Journal of Cognitive Neuroscience*, *24*(6), 1447–1461.

Jäncke, L., Mirzazade, S., & Shah, N. J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neuroscience Letters*, *266*(2), 125–128.

Jenny, J.-P. (2019). Vom Segen des Sezierens. *Retrieved Nov, 13, 2019, from*. https://www.nzz.ch/feuilleton/ueber-das-sezieren-von-leichen-kleine-geschichte-der-anatomie-ld.1469099

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163–178.

Keenan, J. P., Thangaraj, V., Halpern, A. R., & Schlaug, G. (2001). Absolute pitch and planum temporale. *NeuroImage*, *14*(6), 1402–1408.

Kim, S.-G., & Knösche, T. R. (2016). Intracortical myelination in musicians with absolute pitch: Quantitative morphometry using 7-T MRI. *Human Brain Mapping*, *37*(10), 3486–3501.

Kim, S.-G., & Knösche, T. R. (2017). Resting state functional connectivity of the ventral auditory pathway in musicians with absolute pitch. *Human Brain Mapping*, *38*(8), 3899–3916.

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage*, *13*(4), 646–653.

Klein, M. E., & Zatorre, R. J. (2011). A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia*, *49*(5), 878–887.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*(2), 207–212.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*(12), 578–584.

Koole, S. L., & Lakens, D. (2012). Rewarding Replications: A Sure and Simple Way to Improve Psychological Science. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(6), 608–614.

Kriegeskorte, N., Lindquist, M. A., Nichols, T. E., Poldrack, R. A., & Vul, E. (2010). Everything you never wanted to know about circular analysis, but were afraid to ask. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *30*(9), 1551–1557.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Kybic, J., Clerc, M., Abboud, T., Faugeras, O., Keriven, R., & Papadopoulo, T. (2005). A common formalism for the integral formulations of the forward EEG problem. *IEEE Transactions on Medical Imaging*, *24*(1), 12–28.

Lehrl, S., Gallwitz, A., Blaha, L., & Fischer, B. (1992). *Geistige Leistungsfähigkeit - Theorie und Messung der biologischen Intelligenz mit dem Kurztest KAI*. Vless.

Levitin, D. J. (1994). Absolute memory for musical pitch: evidence from the production of learned melodies. *Perception & Psychophysics*, *56*(4), 414–423.

Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: perception, coding, and controversies. *Trends in Cognitive Sciences*, *9*(1), 26–33.

Liem, F., Zaehle, T., Burkhard, A., Jaencke, L., & Meyer, M. (2012). Cortical thickness of supratemporal plane predicts auditory N1 amplitude. *Neuroreport*, *23*(17), 1026–1030.

Li, Y., Ma, Z., Lu, W., & Li, Y. (2006). Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach. *Physiological Measurement*, *27*(4), 425–436.

Loui, P., Li, H. C., Hohmann, A., & Schlaug, G. (2011). Enhanced cortical connectivity in absolute pitch musicians: a model for local hyperconnectivity. *Journal of Cognitive Neuroscience*, *23*(4), 1015–1026.

Loui, P., Zamm, A., & Schlaug, G. (2012). Absolute Pitch and Synesthesia: Two Sides of the Same Coin? Shared and Distinct Neural Substrates of Music Listening. *ICMPC:*

Proceedings / Edited by Catherine Stevens ... [et Al.]. International Conference on Music Perception and Cognition, 618–623.

Loui, P., Zamm, A., & Schlaug, G. (2012). Enhanced functional networks in absolute pitch. *NeuroImage*, *63*(2), 632–640.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luders, E., Gaser, C., Jancke, L., & Schlaug, G. (2004). A voxel-based approach to gray matter asymmetries. In *NeuroImage* (Vol. 22, Issue 2, pp. 656–664). https://doi.org/10.1016/j.neuroimage.2004.01.032

Markham, J. A., & Greenough, W. T. (2004). Experience-driven brain plasticity: beyond the synapse. *Neuron Glia Biology*, *1*(4), 351–363.

McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology*, *18*(4), 452–463.

Meyer, M., Liem, F., Hirsiger, S., Jäncke, L., & Hänggi, J. (2014). Cortical surface area and cortical thickness demonstrate differential structural asymmetry in auditory-related areas of the human cortex. *Cerebral Cortex* , *24*(10), 2541–2552.

Miyazaki, K. 'ichi. (1989). Absolute Pitch Identification: Effects of Timbre and Pitch Region. *Music Perception: An Interdisciplinary Journal*, *7*(1), 1–14.

Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics*, *44*(6), 501–512.

Mori, S., & Zhang, J. (2006). Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*, *51*(5), 527–539.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375–425.

Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jäncke, L., Tempelmann, C., Hinrichs, H., & Heinze, H.-J. (2002). Delayed striate cortical activation during spatial attention. *Neuron*, *35*(3), 575–587.

Oakes, W. F. (1955). An experimental study of pitch naming and pitch discrimination reactions. *The Journal of Genetic Psychology*, *86*(2), 237–259.

Oechslin, M. S., Imfeld, A., Loenneker, T., Meyer, M., & Jäncke, L. (2009). The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. *Frontiers in Human Neuroscience*, *3*, 76.

Oechslin, M. S., Meyer, M., & Jäncke, L. (2010). Absolute pitch--functional evidence of speech-relevant auditory acuity. *Cerebral Cortex* , *20*(2), 447–455.

Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., & Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex* , *11*(8), 754–760.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869.

Ott, C. G. M., Langer, N., Oechslin, M. S., Meyer, M., & Jäncke, L. (2011). Processing of voiced and unvoiced acoustic stimuli in musicians. *Frontiers in Psychology*, *2*, 195.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology*, *94*(1), 26–40.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814.

Petran, L. A. (1932). An experimental study of pitch recognition. In *Psychological Monographs* (Vol. 42, Issue 6, pp. i – 124). https://doi.org/10.1037/h0093299

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time." *Speech Communication*, *41*(1), 245–255.

Profita, J., & Bidder, T. G. (1988). Perfect pitch. *American Journal of Medical Genetics*, *29*(4), 763–771.

Rauschecker, J. P. (2015). Auditory and visual cortex of primates: a comparison of two sensory systems. *The European Journal of Neuroscience*, *41*(5), 579–585.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*(6), 718–724.

Rauscher, F. H., Shaw, G. L., & Ky, C. N. (1993). Music and spatial task performance. In *Nature* (Vol. 365, Issue 6447, pp. 611–611). https://doi.org/10.1038/365611a0

Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science*, *14*(3), 262–266.

Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. *Progress in Brain Research*, *217*, 37–55.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267*(5198), 699–701.

Schmithorst, V. J., & Wilke, M. (2002). Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neuroscience Letters*, *321*(1-2), 57–60.

Schmithorst, V. J., Wilke, M., Dardzinski, B. J., & Holland, S. K. (2002). Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: a cross-sectional diffusion-tensor MR imaging study. *Radiology*, *222*(1), 212–218.

Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., Dosch, H. G., Bleeck, S., Stippich, C., & Rupp, A. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience*, *8*(9), 1241–1247.

Schulze, K., Gaab, N., & Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neuroscience*, *10*, 106.

Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *The European Journal of Neuroscience*, *33*(1), 189–196.

Schulze, K., Mueller, K., & Koelsch, S. (2013). Auditory stroop and absolute pitch: an fMRI study. *Human Brain Mapping*, *34*(7), 1579–1590.

Seldon, H. L. (1981a). Structure of human auditory cortex. I. Cytoarchitectonics and dendritic distributions. In *Brain Research* (Vol. 229, Issue 2, pp. 277–294). https://doi.org/10.1016/0006-8993(81)90994-x

Seldon, H. L. (1981b). Structure of human auditory cortex. II. Axon distributions and morphological correlates of speech perception. In *Brain Research* (Vol. 229, Issue 2, pp. 295–310). https://doi.org/10.1016/0006-8993(81)90995-1

Seldon, H. L. (1982). Structure of human auditory cortex. III. Statistical analysis of dendritic trees. In *Brain Research* (Vol. 249, Issue 2, pp. 211–221). https://doi.org/10.1016/0006-8993(82)90055-5

Seppänen, M., Hämäläinen, J., Pesonen, A.-K., & Tervaniemi, M. (2012). Music training enhances rapid neural plasticity of n1 and p2 source activation for unattended sounds. *Frontiers in Human Neuroscience*, *6*, 43.

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *23*(13), 5545–5552.

Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport*, *16*(16), 1781–1785.

Sinai, A., & Pratt, H. (2003). High-resolution time course of hemispheric dominance revealed by low-resolution electromagnetic tomography. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *114*(7), 1181–1188.

Sirevaag, A. M., & Greenough, W. T. (1987). Differential rearing effects on rat visual cortex synapses. III. Neuronal and glial nuclei, boutons, dendrites, and capillaries. *Brain Research*, *424*(2), 320–332.

Skrandies, W. (1989). Data reduction of multichannel fields: global field power and principal component analysis. *Brain Topography*, *2*(1-2), 73–80.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., & Behrens, T. E. J. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*(4), 1487–1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Michael Brady, J., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219.

Song, J., Davey, C., Poulsen, C., Luu, P., Turovets, S., Anderson, E., Li, K., & Tucker, D. (2015). EEG source localization: Sensor density and head surface coverage. *Journal of Neuroscience Methods*, *256*, 9–21.

Strohmer, J. (2014). *Dorsch-Lexikon der Psychologie*. Wirtz, M.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, *2011*, 879716.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, *113*(2), 345–361.

Terhardt, E., & Dixon Ward, W. (1982). Recognition of musical key: Exploratory study. In *The Journal of the Acoustical Society of America* (Vol. 72, Issue 1, pp. 26–33). https://doi.org/10.1121/1.387989

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and Hearing*, *22*(2), 79–90.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, *13*(7), 967–985.

Van Hecke, W., Emsell, L., & Sunaert, S. (2015). *Diffusion Tensor Imaging: A Practical Handbook*. Springer.

Van Hedger, S. C., Heald, S. L. M., Huang, A., Rutstein, B., & Nusbaum, H. C. (2017). Telling in-tune from out-of-tune: widespread evidence for implicit absolute intonation. *Psychonomic Bulletin & Review*, *24*(2), 481–488.

Vaughan, H. G., Jr, & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalography and Clinical Neurophysiology*, *28*(4), 360–367.

Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.-F., & Pernier, J. (1995). Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis. In *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* (Vol. 96, Issue 2, pp. 143–156). https://doi.org/10.1016/0168-5597(94)00242-7

Vigário, R. N. (1997). Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, *103*(3), 395–404.

Ward, W. D., Dixon Ward, W., & Burns, E. M. (1982). Absolute Pitch. In *Psychology of Music* (pp. 431–451). https://doi.org/10.1016/b978-0-12-213562-0.50018-8

Warren, J. E., Wise, R. J. S., & Warren, J. D. (2005). Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends in Neurosciences*, *28*(12), 636–643.

Wengenroth, M., Blatow, M., Heinecke, A., Reinhardt, J., Stippich, C., Hofmann, E., & Schneider, P. (2014). Increased volume and function of right auditory cortex as a marker for absolute pitch. *Cerebral Cortex* , *24*(5), 1127–1137.

Wernicke, C. (1874). *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*.

Wilson, S. J., Lusher, D., Wan, C. Y., Dudgeon, P., & Reutens, D. C. (2009). The neurocognitive components of pitch processing: insights from absolute pitch. *Cerebral Cortex* , *19*(3), 724–732.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. In *NeuroImage* (Vol. 92, pp. 381–397). https://doi.org/10.1016/j.neuroimage.2014.01.060

Wu, C., Kirk, I. J., Hamm, J. P., & Lim, V. K. (2008). The neural networks involved in pitch labeling of absolute pitch musicians. *Neuroreport*, *19*(8), 851–854.

Wu, W., Wheeler, D. W., Staedtler, E. S., Munk, M. H. J., & Pipa, G. (2008). Behavioral performance modulates spike field coherence in monkey prefrontal cortex. *Neuroreport*, *19*(2), 235–238.

Zaehle, T., Wüstenberg, T., Meyer, M., & Jäncke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. *The European Journal of Neuroscience*, *20*(9), 2447–2456.

Zatorre, R. J. (2003). Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience*, *6*(7), 692–695.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex* , *11*(10), 946–953.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46.

Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., & Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(6), 3172–3177.

Figure 1.1: Broad overview of an ERP paradigm 20
Figure 1.2: Measuring the P2 wave within a predefined time window 23
Figure 1.3: Definition and calculation of area amplitude 25
Figure 1.4: Motion of water molecules 27
Figure 1.5: The concept of a diffusion tensor 28
Figure 2.1: Distribution of AP scores 40
Figure 2.2: Experimental design 42
Figure 2.3: ERPs and topographic scalp maps 45
Figure 2.4: Source estimation 46
Figure 2.5: Time windows used for source reconstruction 47
Figure 2.6: ROIs within each hemisphere 48
Figure 2.7: Signed area values for the components of interest 50
Figure 2.8: Source reconstructions for the N100 and P200 52
Figure 2.9: Correlation between the AP score and the MTG 53
Figure 3.1: Pitch-labeling scores 67
Figure 3.2: Results of the voxel-wise fractional anisotropy analyses 73
Figure 3.3: Correlation between the mean FAs and pitch-labeling score 74

List of Tables

Table 2.1: Demographic variables 38

Table 2.2: Descriptive statistics of demographic variables 39

Table 2.3: Correlations between the AP scores and ERP area 51

Table 3.1: Coordinates of the regions of interest 69-70

Table 3.2: Descriptive statistics 72

Table 3.3: Results of the voxel-wise analysis 75

Table A.1: Results for the whole brain voxelwise analysis 122

A SUPPLEMENTARY MATERIAL

Table A.1: Results for the whole brain voxelwise analysis of the fractional anisotropy data (Study II). The columns indicate the size of the clusters in voxels, the corresponding p-value of the clusters, and the MNI coordinates of the maximal activity. For each contrast, 5,000 random permutations were applied and the significance threshold was set to $\alpha < .20$ (FWE corrected).

No.	Voxel in size	p-value	MNI		
			X	Y	Z
1	138	.177	-39	-40	13
2	85	.176	-26	-3	-13
3	64	.187	25	-31	31
4	31	.195	-34	-32	26

Personal Information

Name	Anja Burkhard
Date of birth	13.04.1985
Nationality	Swiss

Education

2015 – 2019	Ph.D. Student University of Zurich, Switzerland Prof. Dr. rer. nat. Lutz Jäncke Institute of Psychology Division Neuropsychology
2005 – 2014	Lizentiat der Philosophischen Fakultät University of Zurich
1998 – 2004	Gymnasiale Maturität Kantonsschule Alpenquai Luzern

Publications

Liem, F., Zaehle, T., **Burkhard, A.**, Jäncke, L. & Meyer, M. (2012). Cortical thickness of supratemporal plane predicts auditory N1 amplitude. *Neuroreport, 23* (17), 1026 – 1030.

Burkhard, A., Elmer, S., Kara, D., Brauchli, C. & Jäncke, L. (2018). The effect of background music on inhibitory functions: an ERP study. *Frontiers in human neuroscience,* 12, 293.

Brauchli, C., Elmer, S., Rogenmoser, L., **Burkhard, A.**, & Jäncke, L. (2018). Top–down signal transmission and global hyperconnectivity in auditory-visual synesthesia: Evidence from a functional EEG resting-state study. *Human brain mapping*, *39*(1), 522-531.

Jäncke, L., Leipold, S., & **Burkhard, A.** (2018). The neural underpinnings of music listening under different attention conditions. *Neuroreport*, *29*(7), 594-604.

Burkhard, A., Elmer, S. & Jäncke, L. (2019). Early tone categorization in absolute pitch musicians is subserved by the right-sided perisylvian brain. *Scientific reports,* 9(1), 1-14.

Burkhard, A., Hänggi, J., Elmer, S., & Jäncke, L. (2020). The importance of the fibre tracts connecting the planum temporale in absolute pitch possessors. *NeuroImage*, *211*, 116590.